PAUL R. LEPAGE

GOVERNOR

PATRICIA W. AHO

COMMISSIONER

May 5, 2014

Ms. Regina McCarthy
Administrator
Environmental Protection Agency
1200 Pennsylvania Ave., NW
Washington, DC, 20460

RE: Docket ID: **EPA-HQ-OAR-2009-0734**

Proposed Updates to the Residential Wood Heater New Source Performance Standards (NSPS)

Dear Ms. McCarthy,

The Maine Department of Environmental Protection (DEP, "Department") appreciates the Environmental Protection Agency's (EPA) solicitation of comments on the proposed updates to the 1988 Residential Wood Heater New Source Performance Standards (NSPS), 40 CFR Part 60, Subpart AAA, and the proposed new NSPS Subparts QQQQ and RRRR. The Department and the people of Maine take pride in and value that we meet all current National Ambient Air Quality Standards (NAAQS) and appreciate efforts, both large and small, to protect this vital natural resource for our state, region and nation.

Maine supports regulations which result in more efficient and environmentally beneficial wood-burning devices and agrees with the proposal's goal to reduce the amount of wood smoke in our air. Maine also supports emission reductions based on sound, scientifically-proven certification testing methods for such wood burning units. Although the Department applauds and agrees with EPA on several components of the proposal, there are areas of the proposal that create concern for our citizens, our forest products industry, associated manufacturers, and the range of small businesses which exist to meet needs related to residential wood burning in our predominantly rural state.

According to the U.S. Energy Information Administration (EIA), from 2005 – 2012, the northeastern U.S. has seen the highest increase in the use of wood as the main source of household heating.[1] More importantly, in a 2009 study, EIA assessed the percentage of households using wood by income (Figure 1). Clearly, lower income levels have a lower percentage of wood use overall, but a higher percentage

---

[1] U.S. Energy Information Administration. Increase in wood as main source of household heating most notable in the Northeast. March 17, 2014. http://www.eia.gov/todayinenergy/detail.cfm?id=15431

AUGUSTA
17 STATE HOUSE STATION
AUGUSTA, MAINE 04333-0017
(207) 287-7688 FAX: (207) 287-7826

BANGOR
106 HOGAN ROAD, SUITE 6
BANGOR, MAINE 04401
(207) 941-4570 FAX: (207) 941-4584

PORTLAND
312 CANCO ROAD
PORTLAND, MAINE 04103
(207) 822-6300 FAX: (207) 822-6303

PRESQUE ISLE
1235 CENTRAL DRIVE, SKYWAY PARK
PRESQUE ISLE, MAINE 04769
(207) 764-0477 FAX: (207) 760-3143

web site: www.maine.gov/dep

of average wood consumption. This may imply that lower income households are more likely to use wood as their primary heat source as compared to higher income households that use wood as a supplemental heat source. As petroleum-based heating fuel prices have risen in recent years, wood has become more prevalent as both a primary and a secondary heating option, and lower income households have taken advantage of wood as their primary heating method. Figure 1 may also imply that lower income residents are using older, inefficient woodstoves that consume more wood fuel. High costs of purchasing new woodstoves may prohibit many from upgrading to newer, more efficient devices; moreover, the costs of installation, operation, and maintenance must be considered in addition to the initial purchase price of the heating device. In a state that is predominately rural, has the lowest per capita income in New England[2], has some of the oldest housing stock nationally, and is demographically the oldest in the nation[3], Maine citizens should have access to affordable, efficient primary heating sources.   We consider installation costs and considerations, ease of operation, necessary maintenance, and affordability very important factors.   Therefore, we encourage EPA to spearhead energy-savings incentive programs that encourage residents to change out older, less efficient wood stoves for newer, more efficient devices that consume less wood and provide more energy.

Figure 1. Wood use by income. U.S. Energy Information Administration. March 17, 2014.
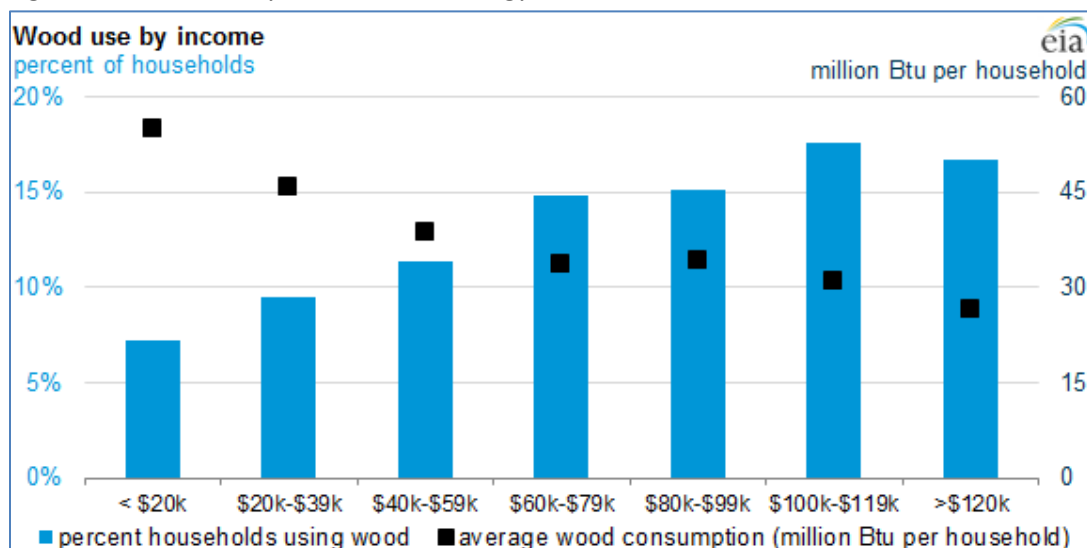
Staff at the Department have worked extensively researching, educating, and reaching out to all parties in the state of Maine that may be affected by EPA's proposal and provide the following comments seeking to improve the proposal.

---

[2] Bureau of Economic Analysis, U.S. Dept. of Commerce: http://www.bea.gov/regional/bearfacts/statebf.cfm
[3] Ranking of States by Projected Percent of Population age 65 and Over: 2000, 2010 and 2030. U.S. Census Bureau: http://www.census.gov/population/www/projections/projectionsagesex.html

We recognize that the industry affected by this regulation is comprised of many small manufacturers who may be without the technology and development resources available to larger manufacturing operations. It is commendable that since promulgation of the 1988 law, these manufacturers have continued, on the whole, to develop and improve their manufacturing processes and products to attain performance levels below the 1988 standards, not because it was mandated, but based on market considerations and judicious business decisions. Maine recognizes the strides taken by this manufacturing sector and thanks them for their efforts and marked progress from which we all benefit.

We agree with the inclusion of currently unregulated units; specifically, the inclusion of single burn-rate stoves, outdoor wood boilers and hydronic heaters, and the wood burning appliances previously exempted from the 1988 NSPS for residential wood burning appliances. We do, however, remain neutral regarding the regulation of masonry heaters in the proposal.

While we strongly support steps to reduce particulate matter and other harmful emissions which can result from burning wood and wood products for heating purposes, Maine DEP is advocating for reasonableness in this regulation such that newer, cleaner wood stoves are a benefit to the consumer – with no compromise to heating ability, no increase in difficulty of operation or on-going maintenance, and which are not prohibitively expensive – so that actual air pollution reductions and health benefits are realized. As appropriate, comments are provided with the specific applicable citation; when applicable to more overarching concepts, they are presented without a specific citation. The Department's comments are organized into the following categories:

I. 40 CFR Part 60, Subpart AAA, *Standards of Performance for New Residential Wood Heaters*
   a. 40 CFR Part 60, § 60.532
   b. 40 CFR Part 60, § 60.533
   c. 40 CFR Part 60, § 60.538
   d. Other Areas of Comment under 40 CFR Part 60, Subpart AAA

II. 40 CFR Part 60, Subpart QQQQ, *Standards of Performance for New Residential Hydronic Heaters and Forced-Air Furnaces*

III. 40 CFR Part 60, Subpart RRRR, *Standards of Performance for New Residential Masonry Heaters*

IV. Proposed Emissions Standards and Test Methods
   a. Regulatory Specifics
   b. Test Method Concerns
   c. Analytical Method Concerns
   d. Test Set Up and Operating Method
   e. Conclusion

V. Conclusions

I.   40 CFR Part 60, Subpart AAA, *Standards of Performance for New Residential Wood Heaters*

   a.   40 CFR Part 60, § 60.532
   What standards and associated requirements must I meet and by when?

   (b) *2015 Particulate Matter Standards.*
   This section of the proposal identifies identical emissions standards (Phase 1 standards) for affected adjustable burn rate wood heater or pellet stoves equipped with catalytic combustors, for affected adjustable burn rate wood heater or pellet stoves not equipped with catalytic combustors, and single burn rate wood heaters.

   The grouping together of all three categories of wood burning appliances (catalyst, non-catalyst, and pellet) is not logical, as the different categories are designed to operate differently and control emissions differently. Additionally, such grouping is not consistent with EPA's practice of establishing other NSPS and National Emission Standard for Hazardous Air Pollutants (NESHAP) standards and requirements. NSPS and NESHAP regulations applicable to boilers, for example, differentiate between various fuels and various combustion unit designs. The major HAP source Boiler MACT regulation specifies 21 different subcategories of boilers and process heaters and specifically identifies standards for each.[4] In accordance with EPA's established precedents, we consider it appropriate to evaluate available data to identify the top performers in each of these three residential wood heater categories and identify emissions standards for each.

   (c) *2020 Particulate Matter Standards.*

   This section of the proposal identifies identical emissions standards (Phase 2 standards) for the three categories of residential wood burning appliances identified above for the 2015 standards.

   We reiterate the inappropriateness of identical emissions standards for units of different technologies and fuels. Beyond that, however, we consider it inappropriate to promulgate Phase 2 standards at all, considering the imprecision and uncertainty of the test methods and resulting data from which the proposed Phase 2 standards were derived. The fuels, test methods, and testing protocols are all highly variable (The Department provides specifics later in these comments.) and do not provide an adequate foundation upon which to base promulgation of Phase 2 standards.

   EPA has promulgated changes to the prescribed compliance test methods, which is essential to obtaining more reliable testing data. However, it is unacceptable to base a proposed standard

---

[4] 40 CFR §63.7499

on data obtained using one set of test methods, and then require compliance to be based on new test methods that have not yet been vetted and verified for veracity and reliability.

Given the inherent variability in the testing processes, we recommend that EPA promulgate the Phase 1 standards and new and/or revised test methods which include changes to the testing processes which reduce the variability inherent in those processes. The promulgation must include a process by which to verify that test results are not within the noise or background of the variability of the test method. Immediately and including Phase 1 compliance certification testing, the evaluation of performance for each unit should be based on multiple tests (a minimum of three) with compliance based on the average of results. Using this new testing approach, sufficient data should be gathered over the proposed five years of Phase 1 applicability to determine what, if any, emissions standards would be appropriate for Phase 2 standards, including a thorough statistical analysis of the foundational data to justify these conclusions. Any such revisions to the process, such as revised test methods and conducting multiple tests upon which to base certification, would require extensive research, data collection, and analysis in order to decisively conclude what emissions level would represent an appropriate Phase 2 standard within the statistically significant range.

Another key revision to the testing process which has been proposed by the EPA is the use of cord wood as the test fuel instead of crib wood (dimensional lumber). EPA has proposed to require testing during the Phase 1 period using both crib wood and cord wood, but only using cord wood for Phase 2 testing. The proposal contains no plan or method to correlate test results using the two different fuels. Crib wood was required as the test fuel in the original residential wood burners NSPS to decrease testing variances due to fuel variability, because crib wood is a more consistent fuel from test to test than cord wood. Each test must be reproducible from lab to lab to within statistically relevant tolerances, yet tests conducted using cord wood, which varies by species, by region, and even from one load to the next, are inherently not as reproducible. This is yet another aspect of the proposed changes which would require extensive research, data collection, and analysis in order to decisively conclude what emissions level would represent an appropriate Phase 2 standard within the statistically significant range.

We do not support the proposal to require appliances to meet a Phase 2 standard in Burn Category 1 and Burn Category 4, when the standard proposed was derived from an average of performance in Burn Categories 1, 2, 3, and 4. There is no data analysis provided which supports the requirement for a unit's emissions from the lowest burn rate and from the maximum burn rate (*not a weighted average*) to comply with the *average* burn rate value of the top 12% for which data has been collected. This is inappropriate and not consistent with the established practice to determine NSPS standards. The average of the four burn categories with

statistically supported upper values for any one burn category would be a much more appropriate and justifiable proposal, assuming that the issue of variability in the test method and corresponding lack of precision in the results could be resolved.

(e) *Pellet Fuel Requirements.*

1. Certification Requirements

Quality standards in the manufacturing and labeling of wood pellets are appropriate. However, we object to the portion of the proposed regulation, as stated on page 6341 of the Federal Register, that "pellets for the certification tests be only those that have been produced under a licensing agreement with the Pellet Fuels Institute, or equivalent…" We do not believe the Pellet Fuels Institute (PFI) certification program should be the only certification option proposed by EPA (with the only other "or equivalent" option left as uncertain and undefined as currently written). To date, the pellet industry has been self-regulating and maintaining high quality standards in response to market demand. We consider that this should be recognized and encouraged, while not legislating control from a non-governmental certifying body made up of potential business competitors.

The Maine DEP has discussed this proposal with pellet manufacturers in Maine and heard that the facilities are not opposed to a fuel pellet specification but want flexibility in how compliance should be met.

We recommend that EPA set the standards and quality compliance expectations, but allow pellet manufacturers flexibility in how to comply. This may include identification of the requirements of a quality management program, citing PFI's program as an example (dated, in case the program is modified in future by PFI and ends up not being what EPA had intended in this rulemaking), but allow each manufacturer the option to develop their own quality assurance program; and EPA should require semi-annual, third-party certification (such as fuel pellets being tested by a reputable laboratory).

We suggest the following language changes in the proposed regulation:
  (e) *Pellet Fuel Requirements*. ~~Operators of~~ Fuels fired in wood heaters that are certified to burn pellet fuels may only burn pellets that ~~have been produced under a licensing agreement with the Pellet Fuels Institute or an equivalent organization approved by the EPA. The pellet fuel must~~ meet the following minimum requirements:
  (1) Density: consistent hardness and energy content with a minimum density of 38 pounds/cubic foot;
  (2) Dimensions: maximum length of 1.5 inches and diameter between 0.230 and 0.285 inches;

(3) Inorganic fines: less than or equal to 1 percent;

(4) Chlorides: less than or equal to 300 parts per million by weight;

(5) Ash content: no more than 2 percent; and

(6) A quality assurance process ~~licensed by~~ similar to the Pellet Fuels Institute's program dated June 1, 2011, or equivalent ~~organization's~~ process; documented and maintained on site and available upon request by the Administrator.

2. Multiple Testing Using a Wide Range of Pellet Grades

The regulation as proposed specifies that once certified, the owners/operators of pellet burning appliances would only be allowed to burn in their appliance the grade of fuel that the appliance manufacturer chose for the appliance certification test and as the manufacturer specifies in the owner's manual. Page 135 of 354 of the proposal states, "Heater manufacturers have indicated to us that market competition will compel them to specify the widest range of grades for which their heaters will properly perform."

This would logically lead to the conclusion that manufacturers of pellet burning appliances will need to test their appliances several times using a wide range of pellet grades. Was this multiple-testing cost factored into the economic analysis?

We proposed the following language change to remedy this potential additional economic burden:

"Once certified, pellet burning appliances would only be allowed to burn the grade of fuel that the appliance manufacturer chose for the appliance certification test or higher grade fuels and the grades of fuel the manufacturer specifies in their owner's manual for the operator to use."

(f) *Prohibited Fuel Types.*

We agree with the identification of prohibited fuels in the regulation, as found in §60.532 (f) *Prohibited Fuel Types*.

(g) *Owner's Manual.* A person must not operate an affected residential wood heater in a manner inconsistent with the owner's manual. The owner's manual must clearly specify that operation in a manner inconsistent with the owner's manual would violate the warranty.

We disagree with the inclusion of such a requirement. This is not enforceable, nor is it appropriate for a NSPS regulation to interfere with the private contract between the owner and the manufacturer.

(h) *Temperature Sensor Requirement.* An affected wood heater equipped with a catalytic combustor must be equipped with a temperature sensor that can monitor combustor gas

stream temperatures within or immediately downstream [within 2.54 centimeters (1 inch)] of the catalytic combustor surface.

This is an inappropriate requirement and should be removed. There is no identified purpose for inclusion of such instrumentation, by whom it would be monitored, what temperature might be appropriate and why, and where the temperature information would be used. There is no point in adding costs when there is no realistic expectation of environmental or health benefit.

b. 40 CFR Part 60, § 60.533
What compliance and certification requirements must I meet and by when?

(b) *Application for Certificate of Compliance.* [The application must include…]
   (6) A copy of the warranties of the model line, including a statement that the warranties are void if the unit is used to burn materials for which the unit is not certified by the EPA.

This does not appear to be enforceable, nor is it appropriate for a NSPS regulation to interfere with a private contract between the owner and the manufacturer.

(m) *Quality Assurance Program.*
   (1) (iii) Within 30 days after approval by the certifying entity, the quality control plan must also be submitted to EPA for review and approval.

In order to provide the manufacturer with a timely response addressing the quality control plan requiring EPA approval, we recommend the following language be added to this section of the regulation:

EPA shall respond within 30 days of submittal of the quality control plan, either that the plan is approved or that the plan is not approved. If a plan is not approved, EPA shall include the reasons for not approving the plan as part of the response. The manufacturer shall have 30 days from the date of receipt of non-approval notification to submit a new version of the quality control plan incorporating corrections for each deficiency as identified by EPA.

(n) *EPA Compliance Audit Testing.*
   (2) (iii) The test must be conducted using the same test method and procedure used to obtain certification or a new test method approved by the EPA Administrator.

   (4) Revocation of Certification.
      (i) If emissions from a wood heater tested under paragraph (n)(2) of this section exceed the certification emission values limit by more than 50 percent, the Administrator will notify the manufacturer that certification for that model line is suspended effective 72 hours from the receipt of the notice, unless the suspension notice is withdrawn by the

Administrator. The suspension will remain in effect until withdrawn by the Administrator, or 30 days from its effective date (if a revocation notice under paragraph (n)(5)(ii) of this section is not issued within that period), or the date of final agency action on revocation, whichever occurs earlier.

(ii)(A) If emissions from a wood heater tested under paragraph (n)(2) of this section exceed the applicable emission limit, the Administrator will notify the manufacturer that certification is revoked for that model line.

Subpart (iii) of § 60.533 (n)(2): It is inappropriate for the EPA Administrator to require – or allow – a "new test method" to be used for audit testing. This would risk introduction of infidelity in the testing methods and resulting test data. The purpose of an audit program is to apply the testing procedure to a sample (selected unit to test) within a population (all manufactured units of the same design and certification as the sample tested) to assure no material exceptions are included in a sampled population.   Ideal outcomes from an audit are test results that corroborate the results of the certification test.   Allowing such an audit test to be conducted using a "new test method" or any method other than that used in initial certification of that product line introduces variability and uncertainty in the usefulness and applicability of the audit results. If introduction of new test methods is necessary, it would be much more appropriately scheduled for testing conducted to renew certification for a particular model line.

Subpart (i) of § 60.533 (n)(4) identifies that if emissions from a wood heater tested according to the applicable paragraph "exceed the certification emission values limit by more than 50 percent," the Administrator will notify the manufacturer of the suspension of the certification for that model line, effective 72 hours from receipt of notice.

Subpart (ii)(A) of § 60.533 (n)(4) identifies that if emissions from a wood heater tested according to the same applicable paragraph "exceed the applicable emission limit," the Administrator will notify the manufacturer of the revocation of certification for that model line.

These two subparts seem either redundant or contradictory to one another. There is no definition provided for "certification emission values limit" from subpart (i), but logic dictates that it could be interpreted to mean one of two things: 1) the applicable emission limit to which the unit is subject, which would make the subpart (ii)(A) redundant; or 2) the emission value identified from the testing done to obtain certification, which allows the possibility of suspension of certification of a unit which actually tests below the applicable emission limit. Hypothetically, a unit whose certification was obtained from testing at 2.0 g/hr could retest at 3.1 g/hr, which would still be under the current limit of 4.5 g/hr. If so, why would it appropriate to suspend certification for that model line? Additionally, the test methods are not sufficiently

accurate at this time to be able to effectively differentiate an emission rate level of 50% under 4.5 grams per hour, which makes this requirement scientifically suspect. Therefore, we question why the subpart (i) cited above is necessary.

The language contained in Subparts (i) and (ii)(A) of § 60.533 (n)(4) should be revised to clearly convey the intent of this part of the regulation.

c. 40 CFR Part 60, § 60.538
What activities are prohibited under this subpart?

(a) No person is permitted to operate an affected wood heater that does not have affixed to it a permanent label pursuant to § 60.536 (b), (c), or (d)(2) through (d)(5).

This particular requirement appears to be unenforceable. One can relate this requirement to that of tags on newly purchased bed mattresses. This similar requirement has only weakened the regulation as it includes requirements that have no reasonable expectation of enforceability. We advocate for inclusion of such activities as "recommended activities" for the owners and operators of residential wood burning appliances.

(c)(1) No commercial owner is permitted to advertise for sale, offer for sale or sell an affected wood heater permanently labeled under § 60.536 (b) or (c) unless:
  (ii) The commercial owner provides any purchaser or transferee with an owner's manual that meets the requirements of § 60.536(f), a copy of the warranty and a moisture meter.

We disagree with the proposal to require direct distribution manufacturers and retailers to provide a moisture meter with the wood heaters at the time of sale, because it is an unnecessary and likely ineffective add-on cost to the sale. There is no way to practically enforce the accurate use of a moisture meter or adherence to recommendations based on what the moisture meter indicates. We advocate for inclusion of such activities as the use of a moisture meter as "recommended activities" for the owners and operators of residential wood burning appliances.

(e) No person is permitted to install or operate an affected wood heater except in a manner consistent with the instructions on its permanent label and in the owner's manual pursuant to § 60.536(f).
(f) No person is permitted to operate an affected wood heater that was originally equipped with a catalytic combustor if the catalytic element is deactivated or removed.
(g) No person is permitted to operate an affected wood heater that has been physically altered to exceed the tolerance limits of its certificate of compliance.

(h) No person is permitted to alter, deface, or remove any permanent label required to be affixed pursuant to § 60.536.

The above subparts (e), (f), (g), and (h) of § 60.538 all appear to be unenforceable. It only weakens a regulation to include requirements with no reasonable expectation of enforceability. We advocate for inclusion of such activities as "recommended activities" for the owners and operators of residential wood burning appliances.

d. Other Areas of Comment under Subpart AAA

1. (page 12 of 354) This portion of the proposal explains that new pellet heater/stove owners and operators would be required to use only the grade of licensed pellet fuels that are included in the heater/stove certification test, or better.

   Due to the lack of enforceability, we are opposed to this requirement for owners and operators. It would be more appropriately included in the owner's manual as a statement of the grade of pellets used in compliance testing and that the manufacturer recommends the use of the same or higher grade of pellets in the appliance for optimum performance.

2. (page 41 of 354) EPA is proposing to require manufacturers to provide warranties on the catalysts and prohibit the operation of catalytic heaters/stoves without a catalyst.

   The manufacturers are not a regulating body; thus, this method to legislate that manufacturers prohibit certain actions by their customer is a misuse of the relationship between manufacturers and their customers. If there were to be consequences of operating an appliance contrary to the owner's manual, those consequences should be per EPA or the delegated authority, not the manufacturer.

3. (page 47 of 354) EPA asks for comments on whether or not to include, in addition to pellet fuel quality assurance requirements, other requirements of best burn practices or adjustments to help insure proper operation, e.g., chimney height and draft specifications, moisture content of wood, and limits on visible emissions.

   We are opposed to the inclusion of best burn practices or adjustments to help insure proper operation, e.g., chimney height and draft specifications, moisture content of wood, and limits on visible emissions. These requirements would be difficult to enforce, and would require a dramatic increase in compliance resources to support such a requirement. Such information would be appropriately included in support information supplied at the time of purchase, but their inclusion in the proposed NSPS is not appropriate.

4. (page 90 of 354) EPA states the following: "We have not determined the potential for consumers to choose other types of fuels and their associated appliances if the consumer costs of wood-fueled appliances increase and at what level that increase would drive consumer choice. Similarly, we have not determined the degree to which better information on the energy of the NSPS appliances will encourage consumers to choose new wood-fueled appliances over other new appliances."

Since costs are a significant part of the justification for this proposal, both health costs savings and increases in manufacturing and selling process of the units, why have these very important factors not been determined and quantified? We are opposed to such an economic experiment, and request that EPA to investigate these areas to better inform how this regulation may affect both affected manufacturers and affected owners/operators.

5. (page 134 of 354) EPA requests specific comments on the need for and level of a possible CO emissions standard and whether CO monitors should be required to help ensure proper operation of the heater and to reduce health and safety concerns for appliances that are installed in occupied areas.

We agree with EPA's conclusion to not require CO testing and reporting, since, according to EPA, the "current data for CO emissions performance and methods of control are not sufficiently robust to support strong CO emission limits…" and that it can be expected that CO emissions will be reduced as a result of the control of PM, because meeting PM standards will be achieved primarily by BSER based on good combustion for those units not utilizing catalysts. We do consider requirement of CO monitors for appliances that are installed in occupied areas to be neither appropriate nor enforceable. In addition, such a requirement is outside of the jurisdiction of NSPS authorization and would be a misuse of the NSPS regulating avenue to mandate such for private residences.

6. (page 137 of 354) Here, EPA states, "Some advocates have suggested that we only allow use of wood certified to a certain moisture level…" We are opposed to this suggestion. Such a requirement could potentially spill over into regulation of every entity in the United States which sells firewood, and we consider this neither the intent nor within the scope of the NSPS program.

7. (page 139 of 354) Six-Month Sell-Through Provision
Wood stove sales are seasonal, and depending on what part of the year in which the six-month period occurs, dealers may likely not be able to sell out the old stock. This could be financially overwhelming for dealers to absorb the cost of leftover, unsold units. We advocate for extending the six-month sell-through provision to one full year for the dealers, and include the six-month provision applicable to manufacturers only.

8. Testing, General Concern

   What would be the woodstove manufacturer's incentive to conduct costly testing on a wide variety of fuels (such as bio-bricks), unless there was a business arrangement – driven by this regulation – between the stove manufacturer and the fuel manufacturer? This regulation as proposed could effectively close out of the market such smaller, innovative-yet-no-more-environmentally-detrimental manufacturers of bio-bricks and other hybrid fuels. We consider this to be an unintended, negative consequence of promulgation as proposed, and support manufacturer-identified appropriate fuels as *recommended* and not *mandated*.

9. Catalyst-Equipped vs. Non-Catalyst-Equipped Wood Burning Appliances

   A. Installation Differences

   According to the network of residential wood heating appliance dealers, installers, and stove and chimney maintenance professionals whom we surveyed, catalyst stoves require more draft than non-catalyst stoves, to overcome the resistance caused by the additional filter (the catalyst) in the exhaust path; thus, a catalyst stove is more draft-sensitive. This could require a homeowner to add onto their chimney to create proper draft for the stove to operate properly. As required by regulation, the owner's manual for a unit must specify the appropriate installation parameters for the stove, including chimney and draft requirements which differ for different wood stove technologies. If a home owner installs a catalyst-equipped wood stove to replace a non-catalyst stove without making adjustments to the chimney, the owner will be violating a federal regulation because of improper installation of a catalyst-equipped wood stove. Manufacturers will have covered themselves legally for these instances, but owners would not. This suggests that an additional requirement must be included for all stoves to be installed by certified installers; this is an additional cost that is not considered in the proposed regulation.

   Hybrid models (consisting of a combination of catalyst and non-catalyst emissions reduction technologies) currently available come with many stove pipe and chimney parameter requirements (chimney height and draft specifications, thimble height from the base of the stove, etc.), with ranges identified within which optimum operation occurs. Given the unique combination of parameter requirements and the multitude of homes with unique design characteristics, installation in an existing home is likely to compromise some installation parameters, keeping a unit from optimal operation. As stated by one chimney sweep we interviewed, "Hybrids can be a serviceman's nightmare."

   B. Cost Differences

   According to EPA table summarizing unit cost impacts (in 2010 $) of the proposal, certified wood heaters have a baseline unit cost of $859. This does not acknowledge the cost differences between catalyst stoves vs. non-catalyst stoves that were reported to us when

we asked several dealers throughout the State of Maine. Replacement catalysts costs, based on an informal survey of Maine woodstove dealers, are between $350 and $500 for the catalyst alone, and an additional $150 to $200 for the service call. Thus, catalyst replacement cost in aggregate is between $500 (best case) and $700 (worst case) at today's prices. This significant added expense for catalyst-equipped stove owners, which will be required several times during the life of the woodstove, provides a substantial economic deterrent to those in the market for a new wood burning, home heating appliance. In fact, many dealers actively discourage customers from catalyst-equipped stoves because of the additional operational requirements and these added costs.

Once a catalyst wears out such that the catalyst is spent and ready for replacement, the stove not only becomes much less efficient as a heat source but also releases much higher levels of pollutants to the atmosphere. At such time, it is quite common for homeowners to bypass the catalyst and continue heating their homes.

C. Operational Differences

According to many residential wood heating appliance owners and operators, as well as dealers, installers, and stove and chimney maintenance professionals whom we surveyed, a major hurdle to overcome for wood burning appliances meeting very low standards – such as catalyst-equipped stoves – is that they only run for such a short period of time (the duration of the burn once the unit is fueled and closed) compared to the older stoves. This is because they (the manufacturer) "plug the firebox with excess air in order to meet the standard." New, compliant designs won't maintain a fire overnight (when operated according to manufacturer's specifications), which means that owners will be less likely to follow the manufacturer's specifications, but will do what they need to do to maintain heat in their homes overnight. In short, these stoves will demonstrate very low emission levels in the lab, but they don't reach those levels in real life. Actual use of catalyst stoves can result in pollution levels greater than EPA-certified non-catalyst stoves because of improper use of the unit (not waiting until flue gases are up to temperature before engaging the catalyst, bypassing the catalyst, not properly maintaining and replacing the catalyst as necessary, etc.).

One owner of two catalyst-equipped stoves shared the following experience: When a catalyst stove is closed most of the way for overnight operation, there's not enough air flow for the exhaust to push through the catalyst, so pressure builds up, resulting in the stove emitting smoke puffs into the room ever ten minutes or so. The homeowner stated, "I've owned two catalyst stoves, and I'll never do that again!"

Stoves equipped with catalysts are perceived by the general populous as cleaner burning than non-catalyst stoves. As a result, people tend to think their stoves and chimneys don't have to be serviced as often, a trend confirmed in our discussions with those in the businesses of wood stove installation and maintenance and chimney sweeping. Not only is this detrimental to the businesses mentioned, but it increases safety risks unnecessarily.

D. Need to Lower Barriers to Replacement Units

Throughout our extensive investigation and research process, the following has been expressed repeatedly by representatives from every sector affected by this proposal: The best way to get the most immediate air quality improvements and reductions in air emissions from residential wood burning is to institute swap-out programs to remove older, dirtier stoves and replace them with newer, cleaner stoves. If new regulations cause stoves to be prohibitively expensive or more likely to be operated improperly (by not replacing a catalyst when it's time, bypassing the catalyst, etc.), or if the new, mandated stove is not as effective a heating source (having to burn more wood to get the same heating value or not being able to keep a fire going overnight), the expected air quality and health benefits will not be realized. A nationwide, older-stove-for-newer-stove replacement program would be the "low-hanging fruit", more easily achievable, in reducing PM emissions from residential wood burning. The more barriers that can be removed from the path of older stove retirement and replacement with newer, cleaner stoves, the greater air quality and related health improvements we can realize, and sooner.

II. 40 CFR Part 60, Subpart QQQQ, *Standards of Performance for New Residential Hydronic Heaters and Forced-Air Furnaces*

The Department supports EPA in the inclusion of new residential hydronic heaters and forced-air furnaces in the NSPS proposal. Since 2008, the State of Maine has regulated outdoor hydronic heaters and recognizes that these units potentially pose an air quality problem if not properly operated. The Department concurs that further emission limit reductions are appropriate at this time, and the addition of forced hot air wood-fired furnaces is also supported.

The Department disagrees with the suggested limitation of hydronic heaters to a defined "heating season," as such units are often used to provide domestic hot water year-round. It is appropriate to identify emissions standards, but it is not appropriate to limit which part of the year the home owner is allowed to use the appliance. It is not within the authority of the NSPS to regulate constraints on when a homeowner heats their home.

Although the Department supports a sell through provision for indoor furnaces, six months is not a reasonable amount of time for this. Wood burning heating appliance sales are seasonal, and

depending on what part of the year in which the six months fall, they may likely not be able to sell out the old stock in time. This would cause significant impact to the dealerships in Maine. In the proposal there is no sell-through provision in Phase 1 for hydronic heaters. The Department believes it is appropriate to include a timeframe for manufacturers to recoup costs on units already manufactured. We advocate for extending the six-month sell-through provision to one full year for the dealers, and include the six-month provision applicable to manufacturers only.

p. 143 of 354: "Manufacturers of hydronic heaters and forced-air furnaces have known for several years that we were drafting this proposal" [pertaining to not providing an allowance transition period for research and development (R&D) to develop new models].

The Department believes that not knowing what the final regulation would look like and contain, it is short-sighted and unfair of the regulator to make such a statement. The manufacturers can surely conduct R&D, but without knowing the target or goal, it is not the most financially responsible expense.

III. **40 CFR Part 60, Subpart RRRR,** *Standards of Performance for New Residential Masonry Heaters*

The Department is neither for nor against the inclusion of masonry heaters in the proposed NSPS update. If EPA moves forward with emission limits for these units, the Department does not support the testing of individual models for compliance. The very nature of residential masonry heaters is that each one is built on-site at a residence, each with its own characteristics and unique features. Also, each masonry heater can be constructed out of various materials, depending on the purchaser's preferences and structural limitations of the site. The suggestion of individual testing of each masonry heater is unacceptable based on the huge expense of such a test and the inappropriateness of conducting such testing at private residences. The Department would recommend that EPA strongly consider approving models based on a computer simulated design.

Masonry heater manufacturers we have consulted expressed that if there is a standard, it would be more appropriate in the units of pounds of pollutant per kilogram of fuel fired, which would take into account the efficiency of the unit and the unique burn profile of a masonry heater. Given the small population of masonry heaters in the United States and the considerable efficiency of such units, we consider that regulation of this source category might be more appropriately handled at the local level, either by states or communities.

IV. **Proposed Emissions Standards and Test Methods for Residential Wood Heaters**

Maine's DEP has researched the EPA-proposed emissions standards for residential wood burning units and the data upon which the proposed standards are based. Our review of the data, the test methods, and the proposed emissions standards has revealed several areas of concern. We

consider that the EPA did not follow the precedents set by the EPA itself in establishing numerical standards for other NSPS and NESHAP (MACT) standards. This aberration from the established procedures does not seem to yield defensible conclusions upon which to base the emission standards as proposed.

a. Regulatory Specifics

The Clean Air Act, reflecting the Clean Air Act Amendments of 1990 (CAAA), specifies the method to identify maximum achievable control technology (MACT) emissions standards under the Air Toxics section (National Emissions Standards for Hazardous Air Pollutants, or NESHAP) as the average emission limitation achieved by the best-performing 12% of existing sources (excluding Lowest Achievable Emission Rate sources); or for categories or subcategories with less than 30 sources, the average emission limitation achieved by the best-performing five sources.[5] The language of Section 111 of the CAA concerning NSPS does not as specifically identify the method to establish numerical standards. However, the EPA has used the MACT method in both NESHAP and NSPS rulemaking since promulgation of the 1990 CAAA (see Cement Kiln NSPS, Major Source Boiler MACT (Boiler MACT), and Commercial and Industrial Solid Waste Incineration (CISWI) regulation preambles and responses to public comments).

b. Test Method Concerns

Understanding of the precision of any test method is extremely important in understanding the limitations of the data generated from the method; and conclusions reached without consideration of these limitations are likely to be misleading and flawed. One element of test method precision is associated with method detection capabilities, and a second is a function of the measurement value. Measurement imprecision is proportionally highest for values measured below or near a particular test method's detection level and proportionally decreases as values increase above the method detection level. As explained in the responses to comments from the CISWI regulation and as reiterated in the responses to comments from the Boiler MACT regulation [FR, Vol. 76, No. 54 / Monday, March 21, 2011 / Rules and Regulations, Major Source Public Comments and Responses, page 15624 and others], EPA should apply the same procedures to the Residential Wood Burning Device NSPS to account for the effects of measurement imprecision associated with a database, including method detection level (MDL) data. [6]

---

[5] Federal Register, Vol. 76, No. 54 / Monday, March 21, 2011 / Rules and Regulations, Major Source Public Comments and Responses, page 15624; and others

[6] The method detection limit (MDL) is defined as the minimum concentration of a substance that can be measured and reported with 99% confidence that the analyte (the substance being tested for; in this case, PM) concentration is greater than zero and is determined from analysis of a sample in a given matrix containing the analyte. [40 CFR, Appendix B to Part 136]

As expressly stated by EPA, the first step is to define an MDL that is representative of the data used in establishing an emissions limit. For CISWI and Boiler MACT, available data was analyzed to identify the highest test-specific MDL. For the Residential Wood Burner NSPS, no such analysis is provided or alluded to. Again, as stated by EPA, the second step in the process is to calculate three times the representative MDL and compare that value to the calculated emissions limit. If three times the MDL is less than the calculated emissions limit, EPA would conclude that measurement variability had been adequately addressed. However, if three times the MDL is greater than the calculated emissions limit, EPA would conclude that the calculated emissions limit did not account for the measurement variability. If this was the case, EPA used the value equal to three times the MDL as the promulgated emissions limit. This procedure ensures measurement variability is adequately addressed in the promulgated emissions limit.

A paper entitled <u>EPA Wood Heater Test Method Variability Study, October 6, 2010</u>, (by Curkeet and Ferguson; incorporated herein by reference; see Appendix A) puts forth the following:

> No rigorous assessment of the precision of the test methods contained in the woodstove NSPS [40 CFR Part 60, Subpart AAA, originally promulgated in 1988 and amended in 1996] and required to be used to certify model lines under that NSPS has ever been performed, despite the fact that the NSPS was promulgated over 25 years ago, more than 700 woodstoves have been tested and certified under it, and a large data base of proficiency test data from accredited laboratories has been assembled.

Please note that there has been no study or statement we are aware of that has refuted this assertion. The paper then proceeds to document the performance of that assessment, using the EPA proficiency test data base and broadly accepted tools for assessing repeatability and reproducibility. Maine DEP staff have not only reviewed this report, but have also conducted independent, in-house assessment of the EPA's proficiency test data base. The remainder of this section describes Maine's findings and concerns.

Maine DEP's conclusions to date support the position that the testing and emissions sampling and quantification methodologies as required by the federal regulation *Standards of Performance for New Residential Wood Heaters*, 40 CFR Part 60, Subpart AAA, are not capable of yielding reproducible results at the limits being advocated in the recently proposed update to 40 CFR Part 60, Subpart AAA, and proposed NSPS Subparts QQQQ and RRRR. Of concern are the lack of reliability and precision (repeatability, reproducibility) of data resulting from the following two sets of procedures required to demonstrate compliance with emission standards for residential wood stoves:

1) Procedures for fueling and operation for emissions testing, including set-up, test fuel properties, test fuel load configuration, loading and start-up time periods, allowable air and fuel adjustments, stove specifications, etc., as provided by Method 28; and

2) The sets of procedures used to sample and quantify emissions as specified in the analytical Methods 5, 5G, 5H, and their equivalents.

c. Analytical Methods

The second set of procedures, the analytical Methods (Method 5 series and their equivalents), have greater conformity than Method 28 to the EPA Precision, Accuracy, Reproducibility, Comparability, and Completeness (PARCC) guidelines. These analytical Methods are well defined with clear procedures; however, it appears they are being used to determine emission concentrations at levels below what they are capable of reliably reporting based on the available Quality Assurance/Quality Control (QA/QC) data. Since its original promulgation, Method 5 has undergone revisions in protocols to improve measurement variability and provide greater precision of reported results. Available data indicates that under ideal conditions, the analytical Methods are capable of a method detection limit (MDL) approaching 1 gram/hour. Accepted good laboratory practice (GLP) protocols for EPA analytical methods use a five- to tenfold buffer to achieve a practical quantitation limit (PQL), the lowest concentration of an analyte (the substance being tested for; in this case, PM) that can be quantified with a suitable statistical degree of confidence. Any value less than the PQL is inappropriate for use to measure compliance with a numerical standard, as its accuracy is considered "estimated", and is only appropriately used for informational purposes. For example, the value may be used to indicate the presence of a substance in conjunction with a complete data set to define the edge parameters.

If a data set exists demonstrating that the analytical Methods can reproducibly generate MDL data at the 0.1 gram/hour emission rate, then the 1 gram/hour value (ten times the MDL value) may be valid. Likewise, if there is round robin data with some type of blind performance data, it would be much easier to evaluate the methods' capabilities and limitations.

Based on the data collected through the EPA required Accredited Laboratory Proficiency Test Program, it appears that the analytical Methods currently in place are neither reliable nor valid below 5 grams/hour at best, and more realistically not valid below 10 grams/hour. In this regard, the Department concurs with the findings documented in the Curkeet and Ferguson paper and comments on the proposed Boiler MACT rule as provided by Mary Sullivan Douglas of NACAA.[7]

---

[7] Comments submitted on behalf of National Association of Clean Air Agencies (NACAA) on the proposed 40 CFR Part 63, Subpart DDDDD; Document Control Number EPA-HQ-OAR-2002-0058-3525-A1, Comment Excerpt Number 11. (See Appendix B)

d. Test Set-Up and Operation Method

Unfortunately, the analytical test methods have much greater definition and reproducibility compared to test Method 28. Method 28 makes a fair attempt to minimize the variables associated with the wood burn test, but the protocols defined in Method 28 do not match real world conditions. Addressing the uncertainty of all of the test method components, evaluation of the propagation of errors associated with the method as presented by Mr. Curkeet consistently shows variably with the coefficient of variation (CV) in excess of 40%. Given that random probability results are those with a CV > 50%, Method 28 gives results that, under ideal and carefully controlled conditions, are not far from random probability results. A more random approach to materials, moisture content, operators, and stoves would generate more realistic variability and have a better relationship to real world conditions. Based on some of the manufacturer supplied data, it does appear that the method can see down to a magnitude of single digit gram/hour outcomes. Experience with this process and the incentive for best case results would again indicate an order of magnitude value would be more in line with actual field tests, i.e., greater than 10 grams/hour.

e. Conclusion

The document *EPA Guidance for Quality Assurance Project Plans, EPA QA/G-5* (EPA/600/R-98/018, February 1998) contains EPA guidelines identifying a detailed process for evaluating both design protocols and data assessment, to be used for all data that may affect regulatory decisions and oversight, and are designed to hold up in court. Appendix H of the document, in particular, defines and addresses data precision, accuracy, representativeness, completeness, and comparability (PARCC) and the implications of each. These are long-established guidelines with references to all aspects of data, including such data as has been collected concerning wood stoves emissions. We believe that rules to be promulgated should also be able to stand up to this rigorous review and that EPA should follow their own QA/QC protocols for understanding the usability of the available data. Application of these protocols would have identified some bounds on how useful the available data is and the limitations of such as the basis of tighter emissions standards.

On pages 146-147 of the proposal, EPA identifies "a key element of the current 1988 NSPS laboratory audit program is the 'round robin' test program"; acknowledges that "we [EPA] have not given this program as much attention as was envisioned in 1988"; proposes greater and more specific adherence to the program; and states "We think these requirements and reminders, combined with the proposed changes in test methods…will help improve inter-laboratory repeatability and reproducibility." It is encouraging that EPA acknowledges that improvement of inter-laboratory repeatability and reproducibility are needed. However,

unquantified gains such as "will help improve" do not provide a sufficient platform from which to promulgate specific numerical emissions standards.

Reduction of air polluting emissions from residential wood burning is an admirable and needed goal; unfortunately, the testing protocols for fuel burning or associated emissions collection and evaluation methods are not capable of reliably reporting values below this magnitude. If errors from both methods are combined, it is arguable that the lowest possible value which could be statistically significant is one no smaller than 30 grams/hour. This is not conducive to reducing emissions, responsibly operating home heating appliances, or reliably demonstrating compliance with either the existing or the proposed emissions standards for residential wood heaters. The test methods must be defined, refined, and improved such that they generate statistically credible results upon which standards and compliance can then be established. Therefore, the state of Maine strongly encourages EPA to refine its compliance test methods to provide future regulatory certainty prior to promulgating standards below the 4.5 gram per hour level.

V. Conclusion

Maine appreciates EPA for moving forward with the proposal and including additional, appropriate categories of wood-fired residential heaters under the NSPS regulatory umbrella. It is our intent that these comments be helpful toward improvements to a more practical and more effective regulation. We support the Phase 1 standards including differentiations between different design technologies; however, until EPA refines its certification test methods to address or account for the associated uncertainties and variables, gathers sufficient data, and puts the data through rigorous analysis and scrutiny, it is inappropriate to promulgate Phase 2 standards that cannot be scientifically justified. Maine supports emission reductions and standards based on sound, scientifically proven certification testing methods for residential wood burning units.

I welcome any questions or comments you may have regarding the Department's submittal.

Sincerely,

/s/ Patricia W. Aho
Patricia W. Aho,
Commissioner
Maine Department of Environmental Protection

CC: Marc Cone, Maine DEP, Bureau of Air Quality

Appendix of Attachments

Appendix A: *EPA Wood Heater Test Method Variability Study, Analysis of Uncertainty, Repeatability and Reproducibility based on the EPA Accredited Laboratory Proficiency Test Database,* by Rick Curkeet, PE, and Robert Ferguson. October 6, 2010.

Appendix B: Comments submitted on behalf of National Association of Clean Air Agencies (NACAA) on the proposed 40 CFR Part 63, Subpart DDDDD; Document Control Number EPA-HQ-OAR-2002-0058-3525-A1, Comment Excerpt Number 11.

EPA Wood Heater Test Method Variability Study

Analysis of Uncertainty, Repeatability and Reproducibility based on the
EPA Accredited Laboratory Proficiency Test Database

Rick Curkeet, PE
Chief Engineer – Hearth Products
Intertek Testing Services

Robert Ferguson
Ferguson, Andors & Company

October 6, 2010

I. Introduction

The precision of any test method is an extremely important tool in understanding the limitations of the data generated from the method. No rigorous assessment of the precision of the test methods contained in the woodstove NSPS and required to be used to certify model lines under that NSPS has ever been performed, despite the fact that the NSPS was promulgated over 20 years ago, more than 700 woodstoves have been tested and certified under it, and a large data base of proficiency test data from accredited laboratories has been assembled. The objective of this paper is to perform that assessment, using the EPA proficiency test data base and broadly accepted tools for assessing repeatability and reproducibility.

The paper consists of four major sections. Section I is this introduction. Section II provides back ground information that will help frame the issues associated with test method precision. Section III is the assessment of test method precision and is comprised of several sub-sections that address the EPA proficiency test data, and analyze some of the sources of variability. The last section (IV) provides conclusions about variability that are supported by the preceding in-depth analyses.

II. Regulatory Background.

The NSPS was proposed on February 18, 1987 (52 Fed. Reg. 4994) and promulgated on February 26, 1988 (53 Fed. Reg. 5860). It is codified at 40 CFR Part 60, Subpart AAA – *Standards of Performance for New Residential Wood Heaters.* The regulation includes the following EPA test methods:

EPA Method 28 – This method prescribes the fueling and operation procedures for emission testing. Among other parameters, it specifies the test fuel properties, test fuel load configuration, pre-test operating conditions, charcoal bed weight at the start of the test run, loading and start-up time, allowable air supply adjustments, fuel adjustments, end of test run determination and allowable heater body temperature differential from start to end of test, and other parameters needed to conduct a wood heater emission test.

EPA Method 5H - This particulate measurement procedure uses an EPA modified Method 5 sampling train which draws a flue gas sample from the wood heater stack and collects particulate in the sampling probe, on a heated 110 mm filter and in a series of ice water chilled impingers. The sampling rate for the sampling train is to be maintained at a constant proportion of the stack gas flow rate and includes a measurement protocol for determining the stack flow at equal time increments. The particulate mass collected in the probe, on the filters, and in the impingers is determined gravimetrically after the removal of uncombined water.

EPA Method 5G-1 – This method and all of the "5G" methods use a dilution tunnel which collects all the effluent from the appliance chimney and draws it, with ambient dilution air, through a straight duct at a uniform velocity. The velocity is high enough to allow accurate flow rate measurement using a Standard or S-Type pitot tube. Method 5G-1 uses a sample train similar to the Method 5 train but

uses two 110 mm filters in series that are unheated and does not involve recovery of material past the second filter. The particulate mass collected in the probe and on the filters is determined gravimetrically after the removal of uncombined water.

EPA Method 5G-2 – This method utilizes the Method 5H sampling train but the sample is taken from the dilution tunnel (as described in Method 5G-1 above) rather than the wood heater stack. Sample recovery is the same as Method 5H.

EPA Method 5G-3 – This method used two 47 mm filters in series and is otherwise similar to Method 5G-1. However the method requires that two identical sample trains be used simultaneously and that the results from each agree within 7.5% of the mean for the result to be valid.

At the time the NSPS was promulgated, there had been no rigorous assessment of the precision of the wood heater test methods. In the preamble accompanying the proposal, EPA had this to say about the inter-laboratory and intra-laboratory precision of Oregon Method 7, which became EPA Method 5H in the regulation:

> [T]he apparent heater-to-heater differences in the data base reflect not only true differences in performance, but also reflect test method precision. Although data are limited, data obtained by Oregon DEQ suggest that the interlab … (sic) four-run weighted average precision at the level of the standards is not greater than ± 1 g/hr.
>
> In contrast, the database upon which the standards are based does not include individual wood heaters tested at more than one laboratory. Therefore, it was agreed that overall and inter-laboratory component of precision should be determined before enforcement tests are performed at laboratories other than the laboratory that initially certified the wood heater. Further, it was agreed that if the overall four-run weighted average precision exceeds ± 1 g/hr, then the interlab component of the precision would be added to the standard when . . . [enforcement] tests are conducted at other than the original certifying laboratory.
>
> The EPA will, by July 1, 1990, either publish in the Federal Register a determination that the inter-laboratory precision cannot be determined, or promulgate revisions reflecting what that precision has been found to be. 52 Fed. Reg. 5010-5011 (February 18, 1987).

As noted above, however, no determination of either intra-lab or inter-lab precision has ever been made.

Section 60.535 of the regulation provides for accreditation for wood heater emission test laboratories. The conditions of accreditation require that both initial (see §60.535(b)(5)) and annual (see §60.535(b)(7)) proficiency testing be conducted by each accredited laboratory. The proficiency test requirements include conducting at least eight test runs (two in each of the burn rate categories described in EPA Method 28) on a wood heater identified (or actually provided) by EPA. The tests are conducted using all EPA wood heater emission measurement methods for which the particular test laboratory is seeking

initial accreditation, or seeking to maintain accreditation. For example, a laboratory that was accredited for Methods 5H and 5G-1 would run both methods during their eight proficiency test runs.

III. <u>Test Method Variability</u>

A. <u>Basic Science</u>

All measurement processes have an inherent element of variability in the measurement result if the process has sufficiently fine resolution. This variability can be evaluated empirically by performing multiple measurements of a single artifact and evaluating the mean and standard deviation of the test results. The relative magnitude of the standard deviation to the mean is a measure of the dispersion (variability) of the measurement process. This measure of dispersion includes variability that originates both from the measurement process and from changes in the artifact being measured, i.e., the inherent variability of the phenomenon being measured. When the measurement involves performance of a process which might be affected by uncontrolled variables or by the ranges allowed for operational parameters within the process, the variability associated with the process is usually much greater than that that arises from the uncertainty of the actual measurements made.

The "rifle" example is a common way of explaining these issues: If a rifle is repeatedly fired at a target, the impacts of variability can be visibly demonstrated. If the rifle is locked into a firing stand and carefully prepared cartridges (bullet plus powder load) are used and if the rounds are fired over a short period of time where the effects of variations in wind, temperature, humidity and barometric pressure conditions are minimized, the cluster of hits on the target should be very tight. The dispersion that is seen can be attributed to the minute differences in the rounds, variations in the rifle barrel as is heats and accumulates residue as well as the other unpreventable small shifts in ambient conditions. If the same experiment is repeated under a wider range of ambient conditions, say on two different days with different wind conditions, one might expect a wider dispersion of results. If the test is repeated with the rifle fired hand-held by the shooter, the dispersion will almost certainly be wider yet. If two different shooters fire the same rifle, the dispersion will widen further. If the shooters use two different rifles of the same model, again the results will almost certainly spread further. And so on and so on. In other words, as the opportunity for variation in the test conditions expands, so does the likelihood that the variation in the results will also increase.

Understanding and quantifying the variability in measurement systems is a scientific discipline. That discipline uses standardized terminology to categorize and organize the various sources of variability. That terminology is as follows:

Any test procedure that results in a numerical measurement of a product or material attribute is subject to a quantifiable value of "**measurement uncertainty**"[1]. Each

---

[1] Two other concepts need brief mention: "accuracy," and "bias". "Accuracy" is the closeness of agreement between a test result and an accepted reference value. Ref.: ASTM E177. "Bias" is the total

measurement that is made as part of the test procedure, and ultimately combined mathematically to produce a result, contributes to the overall uncertainty of that result. The more measurements made, the greater the overall uncertainty.

The closeness of agreement between independent test results obtained under stipulated conditions is the "**precision**" of the measurement. Ref.: ASTM E177. Precision is usually broken down further into:

Repeatability – The closeness of multiple measurements of the same artifact under the same conditions with the same equipment and operator. The symbol "r" is often used to refer to repeatability. Repeatability can be evaluated by replicate tests of the same product at the same laboratory following the defined procedure.

Reproducibility – The closeness of measurements made on the same or presumed to be identical artifacts by different laboratories, equipment and operators. The symbol "R" is often used for reproducibility. Reproducibility can be evaluated for any quantitative measurement system through an Interlaboratory Study or round robin test program.

B. Quantifying Test Method Precision

1. The Available Data

The primary data set available to evaluate test result variability is that generated through the EPA required Accredited Laboratory Proficiency Test Program. These data supplied by EPA provides a substantial data set which allows assessment of the intra-lab repeatability (r), and inter-lab reproducibility (R) for emissions testing using the NSPS specified test methods.

The laboratories that participated in the proficiency testing are:

Apex Environmental	EEMC
EESPC	Engineering Sciences
Intertek-Middleton	Intertek-Montreal
Lokee Testing	Myren Consulting
Northwest Testing	OMNI-Oregon
OMNI-RTP	PFS
Shelton Research	Underwriters Laboratories

Note: For the purposes of this paper, EEMC, EESPC and Lokee have been grouped together as one entity based on the understanding the test equipment remained the same throughout the changes of ownership.

---

systematic error in a measurement, as contrasted to random error. Ref.: ASTM E177. Quantification of both accuracy and bias requires a reference standard or material with known properties and known uncertainty of those properties. Since no wood stove exists that can consistently produce a specific emissions rate, there is no way to evaluate the accuracy or bias of the test procedures. Accordingly, "accuracy" and "bias" will not be discussed further in this paper.

In this program each of these accredited laboratories was supplied with a test appliance which EPA modified in a manner designed to assure that it would be robust enough to operate the same way after being shipped from lab to lab.

Each laboratory was directed to conduct two full series of four test runs (two test runs in each of the four burn rate categories) in each round of proficiency testing. In addition, some laboratories that were accredited for more than one of the EPA methods were required to run both types of sampling equipment during their test runs and submit two sets of data. For example, a lab that was accredited for method 5H and 5G-1 would run both during their eight test runs. All Method 5G-1 and 5G-3 data were "adjusted to Method 5H equivalent[2]" using the formula $E_{Hequiv.} = 1.82 \times E^{0.83}$. Ref.: 40 CFR Part 60, Subpart AAA, Appendix A, Section 6.6, Eq. 5G-5.

In 1987 and 1988, EPA was able to send the first proficiency stove to each accredited laboratory in both years and obtain complete data sets. In 1989 a second proficiency stove was tested at each accredited laboratory, but several laboratories dropped their accreditation in the first two years of the program. From 1990 on, there was a further reduction in the number of laboratories participating and the schedule for proficiency testing was less rigorous. However, one test stove was used from 1993 - 2000. So there are three stoves with multiple test data sets from multiple laboratories. The data from these tests show no pattern of change over time that would indicate deterioration of the test appliance's performance.

Table 1 is a compilation of weighted average emission rates using all eight runs (or more if the lab ran more than 2 in a burn rate category). The weighted average emission rates were computed from individual test run data that were provided to the primary author by EPA upon request in 2006. These data are included in their entirety in Appendix A. The primary author of this paper performed a limited quality check on the data, by comparing a sample of the data to the original test reports submitted by his laboratory and others, and determined that the check sample and the original submissions were identical. Table 1 utilizes all proficiency test program data provided by EPA, and is grouped for each stove tested as well as being categorized by test laboratory and test year.

---

[2] Although this conversion of data can be thought of as adding another layer of uncertainty to the results, especially since the adjustment equation is non-linear and it was itself based on a "best judgment" curve-fit to experimental data (with its own undetermined uncertainty), including the data adjustment is appropriate because it results in an analysis of the precision of the EPA methods as written and used.

Table 1

EPA Proficiency Test Data Summary - EPA Weighted Average Emissions - g/hr

Lab Code	Method	1987	1988	1989	1990	1993	1995	1996	1997	1999	2000	2005
A	5G	2.71										
B	5G	5.96	3.03									
C	5H	2.21	1.51	14.46	6.51							
D	5G	5.72	2.77		7.32	6.24					5.89	15.60
E	5G	6.16	2.69	14.42	13.79		9.40			6.39		10.32
E1	5H	6.09		13.66								
F	5G		4.12	12.53	6.69							
F1	5H			12.40								
G	5G	2.95										
H	5G	19.10										
H1	5H	12.45	6.22	15.39		4.90			4.12	2.88		
I	5G		2.42									
J	5G			17.57		4.86						
K	5G						6.43				13.82	
L	5G							6.06	4.81	4.90		
Stove		Catalytic 1		Non-Cat 1	Non-Cat 2	Non-Cat 3						Non-Cat 4
EPA Cert.		3.1		7.5	4.5	3.6						3.1

2. Data Analysis Options

Since there is no published EPA method for determining test method precision, two separate data analysis methodologies were employed.

- "Macro" Analysis Using Standard Statistical Tools

In Table 2, we present the results of a "macro" analysis of the data using standard statistical tools. Since there is a wide range in the weighted average emissions performance for each stove in the database, both the standard deviation and the coefficient of variation (CV)[3] were computed for each stove to avoid any potentially misleading assessments of variability that might result from the wide range in the calculated means for the individual test stoves.

---

[3] The coefficient of variation represents the ratio of the standard deviation to the mean, and it is a useful statistic for comparing the degree of variation from one data series to another, even if the means are drastically different from each other. The coefficient of variation is useful because the standard deviation of data must always be understood in the context of the mean of the data. The coefficient of variation is a dimensionless number. So when comparing between data sets with different units or widely different means, one should also determine the coefficient of variation for comparison instead of the standard deviation alone. In this case there is a wide range in the average emission performance of the various test stoves and assessing the variability of the results using just the standard deviation without also assessing the CV could be misleading.

Table 2.

EPA Proficiency Test Data Summary - EPA Weighted Average Emissions - g/hr

Lab Code	Method	1987	1988	1989	1990	1993	1995	1996	1997	1999	2000	2005
A	5G	2.71										
B	5G	5.96	3.03									
C	5H	2.21	1.51	14.46	6.51							
D	5G	5.72	2.77		7.32	6.24					5.89	15.60
E	5G	6.16	2.69	14.42	13.79		9.40			6.39		10.32
E1	5H	6.09		13.66								
F	5G		4.12	12.53	6.69							
F1	5H			12.40								
G	5G	2.95										
H	5G	19.10*										
H1	5H	12.45*	6.22	15.39		4.90			4.12	2.88		
I	5G		2.42									
J	5G			17.57		4.86						
K	5G						6.43				13.82	
L	5G							6.06	4.81	4.90		
Stove		Catalytic 1		Non-Cat 1	Non-Cat 2			Non-Cat 3				Non-Cat 4
EPA Cert.		3.1		7.5	4.5			3.6				3.1
Mean		3.90		14.35	8.58			6.21				12.96
SD		1.74		1.78	3.49			2.75				3.73
CV		44.7%		12.4%	40.7%			44.4%				28.8%
N		14 (Lab H/H1 1987 Excluded)		7	4			13				2

* Data exluded as outliers based on Grubb's test for outliers.

From the Table 2 data it is apparent that the repeatability and reproducibility of the test procedures are not very good.

Without applying any statistical tools, one can simply look at the results of testing the same stove in the same laboratory over two or more years to see the range of intra-laboratory results.

And looking at the results from all the laboratories for all the stoves over all the years of the program, it is possible to assess the inter-laboratory precision. Using 2.8 times the standard deviation to estimate the potential range of results that could be expected with a 95% confidence level [Ref.: ASTM E177, 28.1], it is clear that the for any given test series the reproducibility is on the order of $\pm 4.9 - 9.8$ grams per hour.

This means that for any emissions rate measured using the EPA test methods, the result could be 4.9 to 9.8 grams per hour higher or lower if the appliance were tested again at a different laboratory. Even at one standard deviation (68% confidence level), the reproducibility interval is about 1.7 to 3.5 grams per hour, i.e., there is a 32% chance that a new test result would deviate by more than this amount.

The CV calculations further confirm that variability is high (>40 % CV) for the three test stoves with significant data sets but it is not consistent for all units. The 2005 data can be discounted since two data points are too few to draw any valid conclusions. Since CV is calculated from one standard deviation, the implied variability in the actual data is really about 2.8 times the CV at a 95% confidence level. For example, a CV of 40% implies that the results of any one test could vary by +/- 112% from a population mean with a probability of 5% or less of being farther away.

- Precision Analysis Using ASTM E691

Consistent with the National Technology Transfer and Advancement Act, which creates a presumption in favor of using consensus standards, we determined that ASTM E691 – *Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method* was appropriate for use here.

ASTM E691 provides procedures for planning and conducting studies of the precision of test methods, and also provides statistical techniques for performing precision analyses. Since the EPA proficiency test program has been on-going for more than 20 years, the test program design components {Sections 6 – 14) of ASTM E691 are not germane here. However, the analytical procedures can be applied to the EPA proficiency test data for the purposes of determining precision (both r & R). In the following subsections, these procedures are first described generally and then specifically as they were applied to the EPA proficiency test data.

- ASTM E691 Analytical Procedure and Terminology

The analysis utilizes tabular, graphical, and statistical diagnostic tools for evaluating the consistency of the data so that unusual values may be detected and investigated, and also includes the calculation of the numerical measures of precision of the test method pertaining to both within-laboratory repeatability and between-laboratory reproducibility (*r* & *R*).

Table of Test Results—The test results received from the laboratories are arranged in rows and columns. Each column contains the data obtained from all laboratories for one test sample, and each row contains the data from one laboratory for all test samples. The test results, *x*, from one laboratory on one test sample constitute a cell. The results of the following calculations for that test sample are also entered on the table.

Cell Statistics:
Cell Average[4], $\bar{x}$ —This statistic is the cell average for each laboratory.

$$\bar{x} = \sum_{1}^{n} x / n \quad \text{where } n = \text{the number of test results per cell}$$

Average of the Cell Averages, $\bar{\bar{x}}$ —This statistic is the average of all the cell averages for the one material and all laboratories.

$$\bar{\bar{x}} = \sum_{1}^{p} \bar{x} / p \quad \text{where } p = \text{the number of laboratories}$$

Cell Standard Deviation, s —This statistic is the standard deviation of the test results in each cell and is calculated as the square root of the sum of the squares of the cell

---

[4] ASTM E691 refers to the term "average" when defining statistical terminology. "Average" is equivalent to the "arithmetic mean". The term "mean", when used in this paper should be considered interchangeable with "average" as defined in ASTM E691.

value minus the cell average divided by one less than the number of test results in the cell.

$$s = \sqrt{\sum_{1}^{n}(x - \bar{x})^2 /(n-1)}$$

Cell Deviation, d — The cell deviation is calculated by subtracting the average of the cell averages for all laboratories from the cell average for each laboratory.

$$d = \bar{x} - \bar{\bar{x}}$$

Standard Deviation of the Cell Averages, $s_{\bar{x}}$ —This statistic is calculated as the square root of the sum of the squares of the cell deviation divided by one less than the number of laboratories.

$$s_{\bar{x}} = \sqrt{\sum_{1}^{n} d^2 /(p-1)}$$

Precision Statistics
While there are other precision statistics, the fundamental precision statistics of the Interlaboratory Study are the repeatability standard deviation and the reproducibility standard deviation. Other statistics are calculated from these standard deviations.

Repeatability Standard Deviation, s_r—This statistic is calculated as the square root of the sum of the squares of the cell standard deviation divided by the number of laboratories.

$$s_r = \sqrt{\sum_{1}^{p} s^2 / p}$$

Reproducibility Standard Deviation, s_R — This statistic is calculated as the square root of the square of the standard deviation of the cell average plus the square of the repeatability standard deviation (first adjusted for the number of test results) or is equal to s_r, if s_r is larger than s_R.

$$s_R = \sqrt{(s_{\bar{x}})^2 + (s_r)^2 (n-1)/n}$$

Consistency Statistics
The statistical analysis of the data for estimates of the precision statistics is simply a one-way analysis of variance (within- and between-laboratories) carried out separately for each material. Since such an analysis can be invalidated by the presence of severe outliers, it is necessary to first examine the consistency of the data.

Between-laboratory Consistency Statistic, h— This statistic is calculated as the cell deviation divided by the standard deviation of the cell averages.

$$h = d / s_{\bar{x}}$$

Within-laboratory Consistency Statistic, k— This statistic is calculated as the cell standard deviation from one laboratory divided by the repeatability standard deviation of the material.

$$k = s / s_r$$

*Critical Values of the Consistency Statistics—*A table[5] list critical values of the ***h*** and ***k*** consistency statistics at the 0.5 % significance level. The critical values for ***h*** depend on the number of laboratories, ***p***, and the critical values for ***k*** depend both on the number of laboratories, ***p***, and on the number of replicate test results, ***n***, per laboratory per material. When cell values approach or exceed the critical values for ***h*** and ***k***, those cells or laboratories should be investigated for data problems.

- ASTM E691 Precision Analysis

The proficiency test data was then analyzed to determine test method precision using the procedures outlined in ASTM E691 as described above.[6]

Separate tables (Tables 3a-3c) were created for each of the different stoves that were tested over the years the proficiency test program was conducted where there is sufficient data to apply the procedures specified in ASTM E691[7]. Each participating test lab has a row in each table. Since the proficiency test program required two emission tests in each Method 28 burn rate category, it was possible to calculate two EPA weighted average emission results for each proficiency test series. To do this, the first runs conducted in each of the four burn rate categories were grouped together for the purposes of determining the first weighted average emissions. Likewise, the second run in each burn rate category produces the second weighted average emissions result. This methodology was employed since the burn rate category tests were generally not conducted in any specific order and grouping the first runs in each category most closely simulates an actual emission certification test. The numbered columns in each table represent the weighted average emissions results for the grouped data for the different test series on the particular test stove. For example, in Table 3a, Lab A only participated in one year of testing on Catalytic Stove 1. By grouping the eight individual test runs as described above, two weighted average emissions values result for that lab for that stove and are shown in the columns labeled 1 and 2. Lab E participated in both test years and ran two test methods simultaneously in both years. The results from the second method are shown in the row labeled E1. The data grouping methodology results in four separate EPA weighted average emission rates for

---

[5] ASTM E691 Table 5

[6] In some cases, inadequate data precluded using all data in the analyses. In 2005, only two test laboratories conducted proficiency testing and Non-Cat 4 was not tested in any other year. ASTM E691 requires a minimum of three laboratories for a single sample so the 2005 data was not used.

[7] Ideally the data set used for this analysis should contain a minimum of 3 replicate tests from each of at least 6 laboratories. In Tables 3a-3c, only data from the proficiency test stoves that have sufficient data for a statistically meaningful evaluation are presented.

this lab for Catalytic Stove 1 for the 5G method as shown in the columns labeled 1 through 4 and rows labeled E and four additional results for the 5H method in row E1.

The next column shows $\bar{x}$, the mean of all test series results on the tested stove for each lab. In this case, this includes the two test series for each proficiency test and multiple proficiency tests results on the same test stove whenever available. The columns to the right of $\bar{x}$ include the statistical information about the results from each lab for all test series for each sample tested per ASTM E691, as described above. This statistical information includes s, the standard deviation for individual lab results, d, the cell deviation (how much the cell mean deviates from the overall mean), h, a "between laboratory consistency statistic" and k, a "within laboratory consistency statistic". These latter two are measures of whether data included in the analysis should be investigated for problems. Three weighted average emissions data points were excluded from the analysis based on this test. They are marked with * in the tables that follow.   Also included in the table are $\bar{\bar{x}}$, the overall mean for all test series from all labs, $S_{\bar{x}}$, the standard deviation of the overall mean, s_r, the repeatability standard deviation as well as s_R, the reproducibility standard deviation. Finally, values for repeatability (r) and reproducibility (R) are shown to provide the true assessment of the precision demonstrated by the EPA proficiency test data.

Table 3a

										Critical Values	
			x								
Lab	1	2	3	4	$\bar{x}$	s	d	h	k	h-Crit.	k-Crit.
A	2.77	2.67			2.72	0.07	-1.19	-0.95	0.06	2.15	2.36
B	5	6.69	3.44	3.28	4.6	1.59	0.7	0.56	1.27	2.15	1.92
C	2.28	1.78	1.48	1.78	1.83	0.33	-2.08	-1.66	0.26	2.15	1.92
D	5.6	6.09	2.83	2.8	4.33	1.76	0.42	0.34	1.4	2.15	1.92
E	6.53	5.78	4.69	3.4	5.1	1.36	1.19	0.95	1.08	2.15	1.92
E1	7.39	5.02	2.91	2.55	4.47	2.23	0.56	0.45	1.78	2.15	1.92
F	2.81	3.04			2.93	0.16	-0.98	-0.78	0.13	2.15	2.36
G	22.44*	11.19*	5.19	5.37	5.28	0.13	1.37	1.1	0.1	2.15	2.36

1987-1988 Catalytic Stove 1 - ASTM E691 Analysis

$\bar{\bar{x}}$	3.91	
S_x	1.25	95% Confidence (±)
S_r	1.26	Repeatability (r) 3.52 g/hr
S_R	1.62	Reproducibility (R) 4.53 g/hr

* Data excluded as outliers.

Table 3b

1989 Non-Catalytic Stove 1 - ASTM E691 Analysis

Lab	1	2	3	4	$\bar{x}$	s	d	h	k	h-Crit.	k-Crit.
A	13.55	11.65			12.6	1.34	-1.41	-0.1	1.29	2.05	2.30
A1	13.34	10.9			12.12	1.73	-1.89	-0.13	1.65	2.05	2.30
B	13.68	14.03			13.86	0.25	-0.15	-0.01	0.24	2.05	2.30
B1	12.84	13.35			13.1	0.36	-0.91	-0.07	0.35	2.05	2.30
C	14.47	14.31			14.39	0.11	0.38	0.03	0.11	2.05	2.30
D	17.08	16.99			17.04	0.06	3.03	0.22	0.06	2.05	2.30
E	13.81	16.11			14.96	1.63	0.95	0.07	1.56	2.05	2.30

$\bar{\bar{x}}$ = 14.01

s_x = 1.66

s_r = 1.04   Repeatability (r)   2.92   g/hr

s_R = 1.82   Reproducibility (R)   5.1   g/hr

95% Confidence (±)

Table 3c

1993-2000 Non-Catalytic Stove 3 - ASTM E691 Analysis

Lab	1	2	3	4	5	6	$\bar{x}$	s	d	h	k	h-Crit.	k-Crit.
A	10.22	5.93	7.15	5.54	18.52*	7.61	7.29	1.85	-1.41	-0.1	1.29	1.92	1.75
B	7.25	13.06	6.8	8.24			8.84	2.88	-1.89	-0.13	1.65	1.92	1.84
C	5.23	5.2					5.22	0.02	-0.15	-0.01	0.24	1.92	2.22
D	5.61	7.84	3.41	9.25	4.16	5.27	5.92	2.22	-0.91	-0.07	0.35	1.92	1.68
E	7.02	4.15	6.39	3.5	4.08	1.5	4.44	2.01	0.38	0.03	0.11	1.92	1.68
F	7.26	5.56					6.41	1.2	3.03	0.22	0.06	1.92	2.22

$\bar{\bar{x}}$ = 6.35

s_x = 1.56

s_r = 1.92   Repeatability (r)   5.38   g/hr

s_R = 2.28   Reproducibility (R)   6.39   g/hr

95% Confidence (±)

*Data excluded as outlier.

Abbreviations:

s = standard deviation for individual lab results
d = cell deviation = $\bar{x} - \bar{\bar{x}}$
s_x = standard deviation of $\bar{x}$
s_r = repeatability of standard deviation
s_R = reproducibility of standard deviation
h = between laboratory consistency statistic
k = within laboratory consistency statistic

• Conclusions from ASTM E691 Analyses

The Table 3a, 3b and 3c data shows that the repeatability (within lab), *r*, of the weighted average emissions rate determined by the EPA test methods at the 95% confidence level is at best ±2.9 grams per hour and typically about ±3.5 to ±5.4 grams per hour. The reproducibility (between labs), *R*, at the 95% confidence level is ±4.5 to ±6.4 grams per hour.

3. Sources of Variability in EPA Wood Heater Testing

Given that 12 labs (five are still accredited) and all four particulate measurement methods are represented, it is important to try to understand if the repeatability and reproducibility issues are laboratory and method related or arise from a different source – that is, variable performance of the appliance itself.

To evaluate that question, we next investigated the potential sources of variability in the test methods. We started with the emissions measurement methods (5G-1, 5G-2, 5G-3 and 5H), and then addressed Method 28, which specifies how the appliance is to be operated during emission testing.

• Emission Measurement Methods

Method 5G-1

Potential sources of variability in Method 5G-1 include the following:
o Accuracy of dilution tunnel gas flow and sample flow measurement and the resulting consistency of proportionality.
o Recovery of particulates from the probe and front half of filter holder using a solvent cleaner.
o Weighing errors for tare and final weights of filters and evaporation containers used to collect and weigh the catch.

These sources combine to produce an estimated measurement uncertainty of about ±2.7 to 3% of the emission value measured for a typical passing wood stove test. See Appendix C for the derivation of this estimated uncertainty range.

Method 5G-2

Potential sources of variability in Method 5G-2 include the following:
o Accuracy of dilution tunnel gas flow and sample flow measurement and the resulting consistency of proportionality.
o Recovery of particulates from the probe and sampling line using a solvent cleaner.
o Recovery of particulate collected in the impingers and connecting glassware.
o Weighing errors for tare and final weights of filters and evaporation containers used to collect and weigh the catch.

These sources combine to produce an estimated measurement uncertainty of about ±3 – 3.5% of the emission value measured for a typical passing wood stove test. See Appendix C for the derivation of this estimated uncertainty range.

Method 5G-3

Potential sources of variability in Method 5G-3 include the following:
o   Accuracy of dilution tunnel gas flow and sample flow measurement and the resulting consistency of proportionality.
o   Weighing errors for weights of filters, probes including front filter housings and filter seals.

The measurement uncertainty of method 5G-3 has been determined to be approximately ± 2.5% of the emission value measured for a typical passing wood stove test. See Appendix C for the derivation of this estimated uncertainty range.

Method 5H

Potential sources of variability in Method 5H include the following:
•   Accuracy of flue gas flow and sample flow measurement and the resulting consistency of proportionality.
•   Recovery of particulates from the probe and sampling line using a solvent cleaner.
•   Recovery of particulate collected in the impingers and connecting glassware.
•   Weighing errors for tare and final weights of filters and evaporation containers used to collect and weigh the catch.

These sources combine to produce an estimated measurement uncertainty of about ± 20 to 30% of the total particulate mass determination. See Appendix C for the derivation of this estimated uncertainty range.

•   Conclusions Regarding the Contribution of Emissions Measurement Method Uncertainty to Overall Method Variability

It is apparent from the foregoing analysis that the emission measurement methods explain only a small part of the overall variability that has been discussed and quantified in Section III. It is clear, however, from this analysis that some emission measurement methods contribute less to variability than others with Method 5G-3 the best performer in this regard.[8]

---

[8] With the objective of improving the wood heater test methods based on more than twenty years of experience with the current EPA methods, ASTM E06.54 sub-committee was charged with creating new methods. The first step was to specify only one particulate measurement procedure to help reduce the measurement uncertainty associated with multiple measurement methods. This resulted in the subcommittee's development of ASTM E2515, which specifies a procedure that corresponds very closely to EPA Method 5G-3 but has provisions that allow appliance types other than just wood and pellet heaters to be tested. This procedure was selected because it has the smallest measurement uncertainty of the four current EPA methods and contains the dual train measurement which allows for ready detection of invalid results.

- Operating Protocol for the Appliance (EPA Method 28)

  Potential Sources of Variability in Method 28 include the following:
  - Fuel density variation from approximately 30 to 40 lb/ft^3 (dry weight basis).
  - Fuel moisture content variation from 19 to 25% dry basis (varies in uniformity as well as average).
  - Fuel load configuration details.
  - Coal bed size (20-25% of fuel load weight) and pre-burn temperature conditions.
  - Loading time and start-up procedure.
  - Ambient temperature, barometric pressure and humidity.
  - Variations in control settings and resulting burn rates.
  - Random uncontrollable variables such as when and how the fuel load settles, falls and collapses.

Using the EPA proficiency test data, it is possible to examine several of the various Method 28 parameters for their impact on precision.

The first is an analysis of the impacts of burn rate variations. It is known that emissions performance is related to burn rate to some degree with higher burn rates generally producing more complete combustion.

The proficiency test data does not lend itself well to determining burn rate variability in that each lab must adjust air controls to a setting that will produce a burn rate in one of the required categories. The exception is the high burn rate (Category 4) which is always run with the air controls fully open. Table 4 is a compilation of the Category 4 burn rates from the proficiency test program. Note that the specific labs designated by the lab codes were different from year to year.

Table 4

EPA Proficiency Test Data - High Burn Rates - Dry kg/hr

Stove	Year	Lab 1	2	3	4	5	6	7	Mean	SD	CV
Catalytic Stove 1	1987	2.16	2.66	1.74	2.13	2.06	1.69	2.99			
		2.24	2.90	1.89	2.36	2.06	1.74	4.27*			
	1988	1.96	2.08	1.41	2.13	2.27	2.21	1.57			
		2.11	2.13	2.09	2.27	2.92	2.23	2.04	2.15	0.38	17.8%
Non-Cat 1	1989	1.77	2.06	2.50	1.95	1.85					
		1.83	2.32	2.60	2.02	1.91			2.08	0.29	14.0%
Non-Cat 2	1990	4.39	3.05	2.26							
		4.84	3.14	3.49					3.53	0.94	26.7%
Non-Cat 3	1993	3.20	2.91	4.76							
		3.23	3.75	5.18							
	1995	4.08	1.38								
		4.73	1.40								
	1996	3.10									
		3.70									
	1997	3.17	3.24								
		3.29	3.27								
	1999	4.18	3.80	3.73							
		4.49	4.57	4.67							
	2000	4.71	2.90								
		7.09	3.97						3.79	1.15	30.4%
Non-Cat 4	2005	2.54	2.82								
		2.57	3.20						2.78	0.31	11.0%

* Data not included in statistical analysis.

.

The high burn rate variability is substantial, and it is certain that this variability is a contributing factor in the emissions variability, but it is not likely that it is the most important factor. There are many instances in the proficiency test data where two runs in one lab at very similar burn rates produced substantially different emissions rates. See Appendix B for burn rate versus emissions plots of all the individual run data.

Charts 1 and 2 show the relationship between emissions and fuel moisture content and emissions and load weight (assumed to relate to fuel density) for the 121 runs conducted on the catalytic proficiency test stove in 1987 and 1988. Using the "Corner Score" statistical test where a score of >11 indicates potential dependency between variables [Ref.: Mark's Standard Handbook for Mechanical Engineers, 17-22], it is clear that there is no relationship.

Chart 1

1987 &1988 Cat Stove 1 - % M.C. v Emissions

y = -0.1496x + 7.4864
$R^2 = 0.0048$

Corner Score +4

Chart 2

1987 &1988 Cat Stove 1 - Load Wt. v Emissions

y = -0.2276x + 7.368
$R^2 = 0.0048$
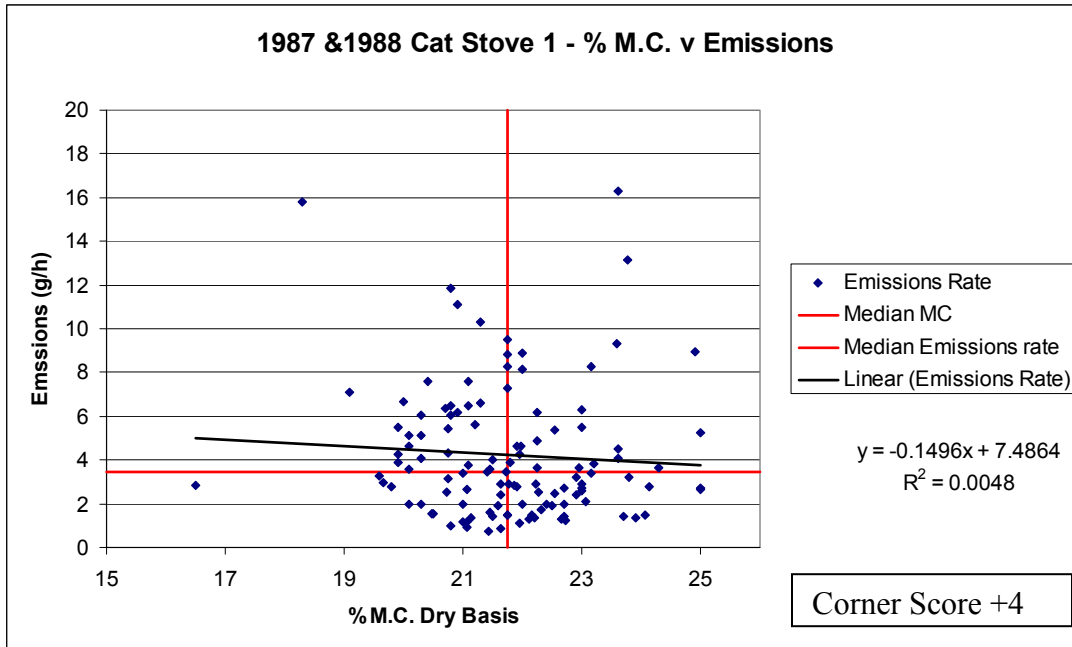
Corner Score +2
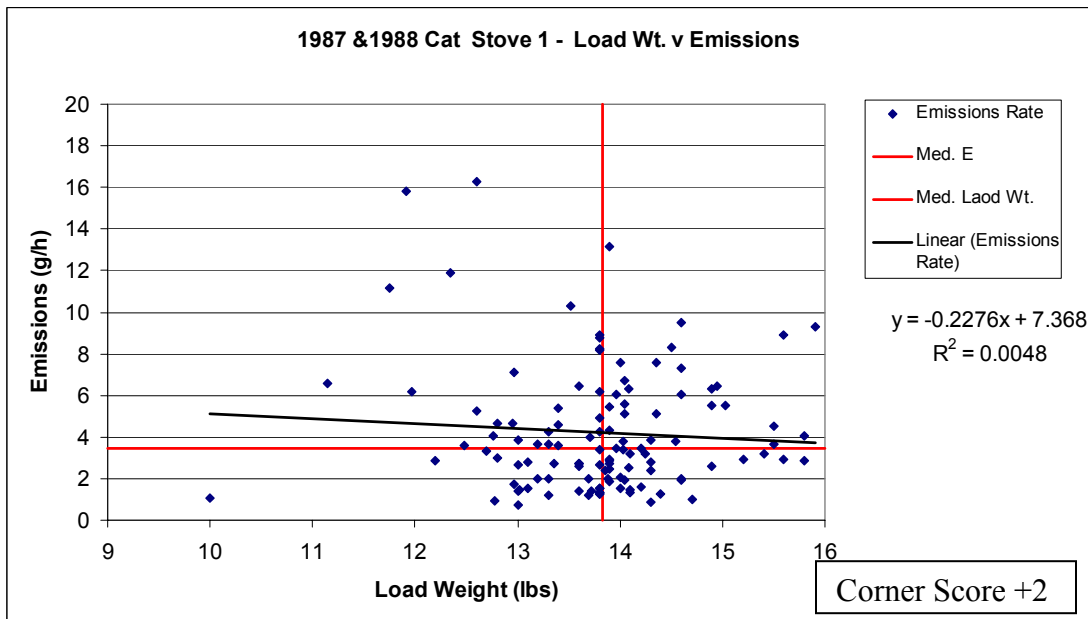
By practical necessity, EPA Method 28 contains tolerances on operational and fueling parameters. These include fuel density, moisture content, laboratory ambient temperature, coal bed size, fuel load weight and several others. While it might be supposed that the method variability could be lessened by even tighter specifications within the test method relating to fuel and operating conditions, the data does not support this proposition.

Without these tolerances or specified ranges wood heater emission testing would become prohibitively expensive with significant additional costs incurred while attempting to control these variables within a significantly narrower range than specified. As demonstrated, these additional costs would not significantly reduce variability.

IV. Final Conclusions

The repeatability and reproducibility of wood heater emission testing as demonstrated by the EPA accredited laboratory proficiency test data is quite poor. At the 95% confidence level, repeatability for the EPA weighted average emission rate is at best ±2.9 g/hour and ranged as high as ±5.4 g/hour. The reproducibility was no better than ±4.5 g/hour and ranged as high as ±6.4 g/hour.

Further analyses of the potential sources of variability have shown that the emission measurement test methods (EPA Methods 5G-1, 5G-2, 5G-3 and 5H) are not major contributors to the high overall variability being demonstrated by the data. However, Method 5G-3 does demonstrate the lowest uncertainty of the four methods currently specified in the NSPS.

Analyses of the variability inherent in EPA Method 28 also indicate that the operational and fuel parameter tolerance ranges specified in that method are not major contributors to the high variability. Tightening these parameters to improve test precision would simply increase costs and the data show that such tightening would not significantly improve precision.

This leaves the only logical conclusion. Variability in wood heater emission testing results for any given appliance is most likely a function of the random nature of burning wood, no matter how tightly you try to control the process. Many relatively small, uncontrollable variables that are inherent in the wood combustion process can combine to significantly affect the outcome of any given test.

These conclusions should not be taken to mean that the current test methodology is inadequate to characterize an appliance's emissions performance. There is no question that the currently certified low emissions wood heaters produce far less particulate emissions than their pre-regulation predecessors which produced emissions rates of 50 to 100 grams per hour or more. Our primary conclusion is that the current testing process simply cannot consistently distinguish emissions performance differences of less than 3 to 6 grams per hour. The process is certainly capable of reliably distinguishing between good and bad performance, but it cannot reliably distinguish between "good, better and best" performance.[9]

---

[9] To return to the "rifle" analogy, wood stoves are not modern high powered rifles in the hands of an expert marksman who can place a dozen shots in a 2 inch bull's-eye at 100 yards. They are rather more like a smooth bore flintlock pistol where just "hitting the paper" at 50 feet is an accomplishment.

Appendix A. – Raw Data

Note: The laboratory codes indicated in this appendix are not those used in the tables in the body of the paper since EPA assigned the lab codes to different labs from year to year. Labs have not been identified by name to maintain the anonymity of the participants.

Year	**1987**				
Stove	Catalytic -1				

Lab A	**Run**	Burn Rate (Kg/h)	Test Load Wt. (lb)	% M **WET**	Emissions (g/h)
5G	1	1.15	13.90	18.40	2.48
	2	2.16	13.30	18.00	4.25
	3	1.15	13.00	17.40	2.69
	4	1.35	13.30	18.20	3.67
	5	2.24	13.80	16.60	4.25
	6	0.79	13.30	17.40	1.18
	7	1.74	13.00	16.60	3.88
	8	0.73	10.00	18.00	1.09

Lab B	**Run**			**DRY**	
5G	0	1.10	12.69	19.60	3.30
	1	2.90	12.76	18.90	28.11
	2	2.16	11.92	18.30	15.79
	3	0.56	12.19	16.50	2.83
	4	1.10	12.96	19.10	7.10
	5	0.99	12.76	20.30	4.07
	6	2.66	11.75	20.90	11.13
	7	2.03	11.97	20.90	6.18
	8	2.10	12.34	20.80	11.87
	9	1.69	11.15	21.30	6.60

Lab C	**Run**			**DRY**	
5H	1	1.42	14.10	22.90	3.20
	2	1.89	13.20	24.29	3.63
	3	1.74	14.30	24.13	2.79
	4	1.71	14.00	23.07	2.08
	5	1.12	14.60	21.60	1.91
	6	1.21	13.80	22.20	1.35
	7	0.87	13.80	22.67	1.27
	8	0.97	14.30	21.63	0.89
	9	0.61	15.40	20.75	3.16

Lab D	Run			DRY	
5G	1	2.13	15.03	19.90	5.50
	2	2.36	14.05	20.00	6.68
	3	1.36	13.97	20.30	6.07
	4	0.83	14.05	21.20	5.60
	5	0.57	14.03	21.00	3.40
	6	1.43	14.01	21.08	7.60
	7	1.10	14.05	20.10	5.14
	8	1.06	14.08	20.70	6.34

Lab E	Run			WET	
5G	1	2.06	14.90	18.70	6.30
	2	2.06	13.80	18.03	8.15
	3	1.56	14.60	17.86	7.30
	4	0.67	13.80	17.86	1.50
	5	1.34	13.80	18.20	6.15
	6	0.59	13.70	17.36	2.00
	7	0.91	13.80	17.86	8.25
	8	0.83	13.90	17.18	5.45

Lab E1	Run			WET	
5H	1	2.06	14.90	18.70	5.50
	2	2.06	13.80	18.03	8.90
	3	1.56	14.60	17.86	9.50
	4	0.67	13.80	17.86	1.50
	5	1.34	13.80	18.20	4.90
	6	0.59	13.70	17.36	1.20
	7	0.91	13.80	17.86	8.80
	8	0.83	13.90	17.18	4.35

Lab F	Run			DRY	
5G	1	1.69	13.52	21.30	10.30
	2	1.63	14.03	23.20	3.80
	3	0.99	13.36	23.00	2.70
	4	0.93	13.86	22.90	2.40
	5	0.93	13.88	22.70	2.00
	6	0.49	13.72	23.70	1.40
	7	1.74	14.25	23.80	3.20
	8	1.62	13.71	21.50	4.00

Lab G	Run			WET	
5G	1	4.27	14.50	18.80	8.29
	2	1.02	14.10	19.20	35.98
	3	0.72	14.30	19.00	19.49
	4	0.82	13.90	19.10	17.01
	5	0.77	14.10	19.30	1.36
	6	1.53	12.60	19.10	16.29
	7	1.46	13.90	19.20	13.16
	8	2.99	13.90	18.70	2.91

Lab H	Run				
5H	1	1.68	n/a	n/a	5.537
	2	1.12			5.713
	3	1.72			6.62
	4	1.01			7.531
	5	2.22			8.08
	6	2.24			19.018
	7	0.7			7.213
	8	0.44			2.386

Year **1988**

Stove Catalytic-1

Lab A	Run	Burn Rate (Kg/h)	Test Load Wt. (lb)	% M WET	Emissions (g/h)
5H	1	1.51	15.20	17.88	2.91
	2	1.96	15.50	18.67	3.66
	3	0.60	15.60	18.19	2.89
	4	0.65	15.90	19.09	9.32
	5	1.45	15.50	19.11	4.49
	6	1.20	15.60	19.94	8.94
	7	2.11	15.80	19.11	4.08
	8	1.22	15.80	17.94	2.85

Lab B	Run			WET	
5G	1	1.47	14.90	18.70	2.62
	2	1.03	13.60	20.00	2.72
	3	1.92	13.80	20.00	2.63
	4	2.08	12.60	20.00	5.25
	5	0.64	14.10	19.40	1.47

	6	0.61	13.10	17.00	1.54
	7	1.03	13.20	18.30	1.98
	8	2.13	13.40	18.40	5.38

Lab C	Run			DRY	
5G	1	1.11	13.90	21.63	2.88
	2	0.87	13.60	20.73	2.56
	3	1.41	13.80	23.17	3.38
	4	0.75	13.80	21.13	1.34
	5	2.12	12.80	20.08	4.66
	6	0.59	13.00	22.71	1.40
	7	2.09	12.80	19.67	2.96
	8	1.32	13.10	21.92	2.80

Lab D	Run			DRY	
5H	1	0.45	13.00	21.43	0.74
	2	2.13	14.30	21.64	2.39
	3	2.27	14.20	21.42	3.45
	4	1.82	13.90	22.50	1.89
	5	1.38	13.60	21.50	1.39
	6	0.76	14.40	22.73	1.23
	7	0.90	13.80	22.12	1.29
	8	1.18	14.20	21.46	1.61

Lab E	Run			DRY	
5G	1	1.82	13.97	21.73	3.48
	2	2.92	12.96	22.32	1.71
	3	0.65	13.02	22.17	1.47
	4	1.83	12.95	21.97	4.64
	5	1.18	12.48	21.46	3.57
	6	1.14	14.09	22.28	2.55
	7	0.47	12.77	21.06	0.93
	8	2.27	12.62	22.10	5.06

Lab F	Run			DRY	
5G	1	2.23	14.35	20.30	5.13
	2	1.85	14.95	20.80	6.47
	3	0.96	14.60	20.10	1.98
	4	1.47	14.55	21.10	3.77
	5	2.21	14.35	20.40	7.58

		6	0.49	14.70	20.80	0.99
		7	0.76	14.00	20.50	1.52
		8	1.03	14.05	20.30	1.96

Lab G	**Run**				DRY	
5G	1	1.43	14.60	20.80	6.03
	2	2.04	13.60	21.10	6.47
	3	1.03	14.30	21.80	3.86
	4	0.75	13.90	22.70	2.73
	5	0.71	13.30	22.00	1.98
	6	0.93	13.60	19.80	2.75
	7	1.57	13.40	21.90	4.60
	8	1.17	13.40	20.10	3.61

Year **1989**

Stove Non-Catalytic-1

Lab A	**Run**	Burn Rate (Kg/h)	Test Load Wt. (lb)	% M **DRY**	Emissions (g/h)
5G	1	1.53	11.50	21.98	15.80
	2	0.89	9.80	22.38	10.95
	3	1.18	10.00	20.33	9.70
	4	0.58	10.40	20.56	20.10
	5	1.83	10.00	20.83	6.60
	6	0.91	9.80	20.56	11.35
	7	0.62	9.70	20.44	18.00
	8	1.77	10.40	20.46	9.85

Lab A1	**Run**			**DRY**	
5H	1	1.53	11.50	21.98	16.10
	2	0.89	9.80	22.38	9.80
	3	1.18	10.00	20.33	9.35
	4	0.58	10.40	20.56	21.00
	5	1.83	10.00	20.83	9.85
	6	0.91	9.80	20.56	10.25
	7	0.62	9.70	20.44	17.25
	8	1.77	10.40	20.46	9.10

Lab B	**Run**			**DRY**	
5G	1	0.80	10.90	21.63	22.18
	2	1.04	11.20	20.85	13.82

	3	0.48	11.20	20.19	18.07
	4	0.98	11.40	21.98	14.19
	5	1.29	10.70	19.93	11.71
	6	1.58	11.40	23.48	13.69
	7	2.32	11.50	22.47	7.55
	8	2.06	11.60	23.05	7.46

Lab B1	**Run**			**DRY**	
5G	1	0.80	10.90	21.63	22.05
	2	1.04	11.20	20.85	12.04
	3	0.48	11.20	20.19	18.39
	4	0.98	11.40	21.98	13.60
	5	1.29	10.70	19.93	10.93
	6	1.58	11.40	23.48	12.24
	7	2.32	11.50	22.47	6.83
	8	2.06	11.60	23.05	6.80

Lab C	**Run**			**DRY**	
5H	1	0.51	10.90	22.71	21.08
	2	0.79	10.90	22.15	14.41
	3	1.03	11.50	21.83	15.55
	4	0.92	11.50	21.66	16.28
	5	1.33	11.20	25.13	12.69
	6	1.47	11.20	21.88	15.05
	7	2.50	11.10	21.13	8.28
	8	2.60	11.60	21.73	8.54

Lab D	**Run**			**DRY**	
5G	1	1.53	10.90	21.88	17.75
	2	1.39	11.10	23.82	11.20
	3	1.95	11.40	23.34	14.02
	4	0.64	11.30	21.28	23.31
	5	1.03	11.30	21.74	19.69
	6	0.79	11.40	21.16	18.38
	7	2.02	10.50	21.75	4.93
	8	1.24	10.40	23.17	20.14

Lab E	**Run**			**DRY**	
5H	1	INVALID			
	2	1.7112	11.5	18.223	5.413

	Run				
	3	0.8368	11.5	18.534	15.492
	4	1.3117	11.5	18.351	19.298
	5	1.8494	11.6	18.066	4.558
	6	1.1465	11.3	18.066	20.368
	7	1.1703	10.5	18.172	11.303
	8	0.8284	11.5	18.027	22.939
	9	1.9069	9.7	17.012	8.424

Year **1990**

Stove Non-Catalytic-2

Lab A	**Run**	Burn Rate (Kg/h)	Test Load Wt. (lb)	% M **DRY**	Emissions (g/h)
5H	1	1.70	17.69	21.49	3.07
	2	1.54	16.84	20.25	2.71
	3	1.02	17.36	21.00	5.68
	4	0.63	18.35	22.48	24.09
	5	4.39	19.03	23.51	8.92
	6	1.10	17.50	21.21	4.91
	7	4.84	18.58	22.82	11.23
	8	0.78	18.01	21.97	2.97

Lab B	**Run**			**DRY**	
5G	1	1.91	10.40	22.34	17.40
	2	0.64	9.50	22.33	8.80
	3	1.29	9.50	22.81	7.82
	4	1.09	10.00	22.33	5.04
	5	3.14	9.90	22.60	12.47
	6	0.99	9.60	23.48	7.82
	7	3.05	9.50	22.77	11.98
	8	0.74	10.70	21.60	38.04

Lab C	**Run**			**DRY**	
5G	1	2.18	10.40	22.63	13.36
	2	0.67	9.90	20.37	5.14
	3	3.49	10.80	21.95	14.89
	4	1.24	10.30	22.00	2.64
	5	1.24	10.40	21.23	2.63
	6	2.26	10.60	21.68	6.27
	7	1.28	9.80	22.23	5.66
	8	0.82	9.50	21.23	5.52

Lab D	Run			DRY	
5G3	1	4.375	9.36	21.33	10.26
	2	0.708	10.41	20.59	6.28
	3	0.639	10.93	19.89	7.24
	4	1.675	9.64	20.42	6.90
	5	1.665	10.45	20.42	9.58
	6	1.478	10.65	21.25	6.90
	7	1.387	10.34	21.65	7.50
	8	4.147	10.69	23.23	9.73

Year **1993**

Stove Non-Catalytic-3

Lab A	Run	Burn Rate (Kg/h)	Test Load Wt. (lb)	% M WET	Emissions (g/h)
5H	1	0.83	10.10	18.06	10.98
	2	3.23	10.10	18.06	13.02
	3	1.22	10.40	18.41	3.34
	4	1.41	10.50	18.39	1.62
	5	1.74	10.60	18.01	2.32
	6	3.20	10.00	17.69	8.20
	7	1.14	10.20	17.93	2.90
	8	0.88	10.20	17.56	6.30

Lab B	Run			DRY	
5G	1	3.75	14.40	23.65	9.31
	2	0.67	10.20	23.20	7.13
	3	2.91	10.30	22.57	5.12
	4	0.81	11.60	22.73	7.79
	5	1.20	10.90	22.48	3.72
	6	1.20	9.90	23.20	5.16
	7	1.36	12.30	21.87	1.54
	8	1.29	12.50	22.23	2.01

Lab C	Run			WET	
5G	1	0.94	9.56	18.82	7.02
	2	0.95	9.64	18.91	7.20
	3	1.07	10.17	19.10	5.79
	4	1.18	9.89	19.61	5.28
	5	1.39	10.04	18.81	5.35
	6	1.47	10.17	17.98	1.66

7	5.18	9.99	18.17	28.47
8	4.76	10.38	19.13	11.89

Year **1995**
Stove Non-Catalytic-3

Lab A	Run	Burn Rate (Kg/h)	Test Load Wt. (lb)	% M WET	Emissions (g/h)
5G	1	1.24	10.10	16.76	3.70
	2	1.07	10.20	18.82	12.00
	3	1.70	10.20	18.52	2.85
	4	4.08	11.10	19.03	11.26
	5	1.20	11.40	19.76	19.90
	6	4.73	11.60	17.57	14.44
	7	1.51	11.50	16.06	8.29
	8	0.95	10.50	18.69	11.00

Lab B	Run			DRY	
5G	1	0.70	10.29	21.90	14.46
	2	0.74	9.66	19.80	10.64
	3	0.87	10.83	19.90	7.49
	4	1.12	9.55	18.50	6.17
	5	1.38	9.80	19.10	4.26
	6	1.40	10.70	20.10	1.26
	7				
	8				

Year **1996**
Stove Non-Catalytic-3

Lab A	Run	Burn Rate (Kg/h)	Test Load Wt. (lb)	% M WET	Emissions (g/h)
5G	1	1.24	9.70	17.79	4.12
	2	0.94	9.30	18.04	9.45
	3	3.10	9.80	18.56	4.02
	4	1.59	10.30	17.96	2.84
	5	0.93	9.90	17.56	10.19
	6	3.70	9.80	16.81	11.27
	7	1.36	10.20	16.84	3.22
	8	1.13	10.60	17.79	8.07

Year **1997**
Stove Non-Catalytic-3

Lab A	Run	Burn Rate (Kg/h)	Test Load Wt. (lb)	% M WET	Emissions (g/h)
5G	1	0.94	10.00	17.48	3.59
	2	3.29	10.70	17.46	3.01
	3	1.42	10.10	17.65	4.96
	4	0.98	10.70	17.76	6.62
	5	1.13	10.60	17.84	4.70
	6	1.16	10.80	17.51	3.58
	7	3.17	10.30	14.51	6.66
	8	1.29	11.00	17.76	3.72

Lab B	Run			DRY	
5H	1	0.85	10.70	16.96	12.56
	2	0.79	10.20	17.62	3.78
	3	3.27	10.40	19.02	8.85
	4	1.22	10.60	17.75	0.82
	5	3.24	10.80	17.32	5.21
	6	1.32	10.50	16.96	1.18
	7	1.46	11.20	18.45	2.45
	8	1.06	11.10	17.75	3.54

Year **1999**
Stove Non-Catalytic-3

Lab A	Run	Burn Rate (Kg/h)	Test Load Wt. (lb)	% M WET	Emissions (g/h)
5H	1	1.24	10.80	17.84	2.50
	2	4.18	11.30	19.39	5.53
	3	1.44	10.00	17.89	3.09
	4	4.49	10.20	19.10	4.78
	5	1.20	10.00	18.23	0.87
	6	1.61	10.20	19.02	0.96
	7	0.91	9.80	16.41	5.24
	8	0.98	10.00	17.27	1.28

Lab B	Run			DRY	
5G	1	1.16	10.60	23.87	6.06

	Run	Burn Rate (Kg/h)	Test Load Wt. (lb)	% M	Emissions (g/h)
	2	3.80	10.40	24.25	9.26
	3	1.62	10.30	23.71	4.11
	4	1.10	10.50	23.14	6.74
	5	0.99	10.20	21.31	9.16
	6	4.57	10.20	21.55	15.99
	7	1.32	10.60	21.68	3.34
	8	0.97	9.70	23.41	8.72

Lab C	**Run**			**WET**	
5G	1	0.966	10.50	17.18	2.42
	2	3.836	9.30	17.44	5.27
	3	1.150	9.50	17.69	3.76
	4	4.672	9.60	17.75	17.97
	5	0.617	9.70	17.70	4.35
	6	0.814	9.60	17.41	12.08
	7	1.262	9.40	17.63	2.45
	8	1.268	9.70	17.37	3.32

Year **2000**

Stove Non-Catalytic-3

Lab A	**Run**	Burn Rate (Kg/h)	Test Load Wt. (lb)	% M DRY	Emissions (g/h)
5G	1	1.26	10.08	22.51	6.10
	2	0.91	10.16	20.97	5.62
	3	0.79	10.04	21.65	7.42
	4	1.17	10.26	20.36	5.34
	5	1.11	10.19	23.05	4.72
	6	1.43	10.19	22.47	1.74
	7	4.71	10.16	22.37	11.40
	8	7.09	10.31	23.68	11.06

Lab B	**Run**			**DRY**	
5G	1	1.287	10.88	20.20	4.64
	2	0.699	10.33	19.70	59.60
	3	0.983	11.05	20.80	7.09
	4	3.967	10.76	21.20	19.82
	5	0.955	9.78	20.80	7.54
	6	1.428	9.62	21.30	4.74
	7	0.803	9.85	21.80	7.75
	8	2.901	10.67	22.00	14.42

Year **2005**

Stove Non-Catalytic-4

Lab A	Run	Burn Rate (Kg/h)	Test Load Wt. (lb)	% M DRY	Emissions (g/h)
5G	1	2.54	9.10	22.00	6.30
	2	1.33	8.20	19.50	11.30
	3	0.54	9.00	19.20	24.30
	4	0.95	9.20	19.40	9.50
	5	0.91	8.40	20.10	10.80
	6	0.70	8.40	21.50	13.20
	7	1.36	8.00	23.20	5.60
	8	2.57	9.10	20.30	7.40

Lab B	Run			DRY	
5G	1	INVALID	8.89	21.79	
	2	0.91	8.34	20.85	23.82
	3	1.55	8.23	20.76	14.59
	4	3.20	8.13	21.51	6.10
	5	1.31	8.05	19.74	9.94
	6	0.55	8.05	20.44	19.55
	7	1.23	8.22	20.87	10.04
	8	2.83	8.23	20.14	7.93

Appendix B – Burn Rate versus Emissions Rate Charts

Emissions v. Burn Rate 1987-1988
Cat Stove 1

$y = 2.4542x + 2.9822$
$R^2 = 0.0794$

$y = 1.3643x + 1.4338$
$R^2 = 0.1816$

Burn Rate V. Emissions - 1989
Non-Cat Stove 1

$y = -6.689x + 22.299$
$R^2 = 0.5151$

Burn Rate v. Emissions
Non-Cat Stove 2

$y = 0.5034x + 8.2554$
$R^2 = 0.008$

Burn Rate v. Emissions
Non-Cat Stove 3

Burn Rate v. Emissions
Non-Cat Stove 4

$y = -4.6817x + 19.037$
$R^2 = 0.4428$
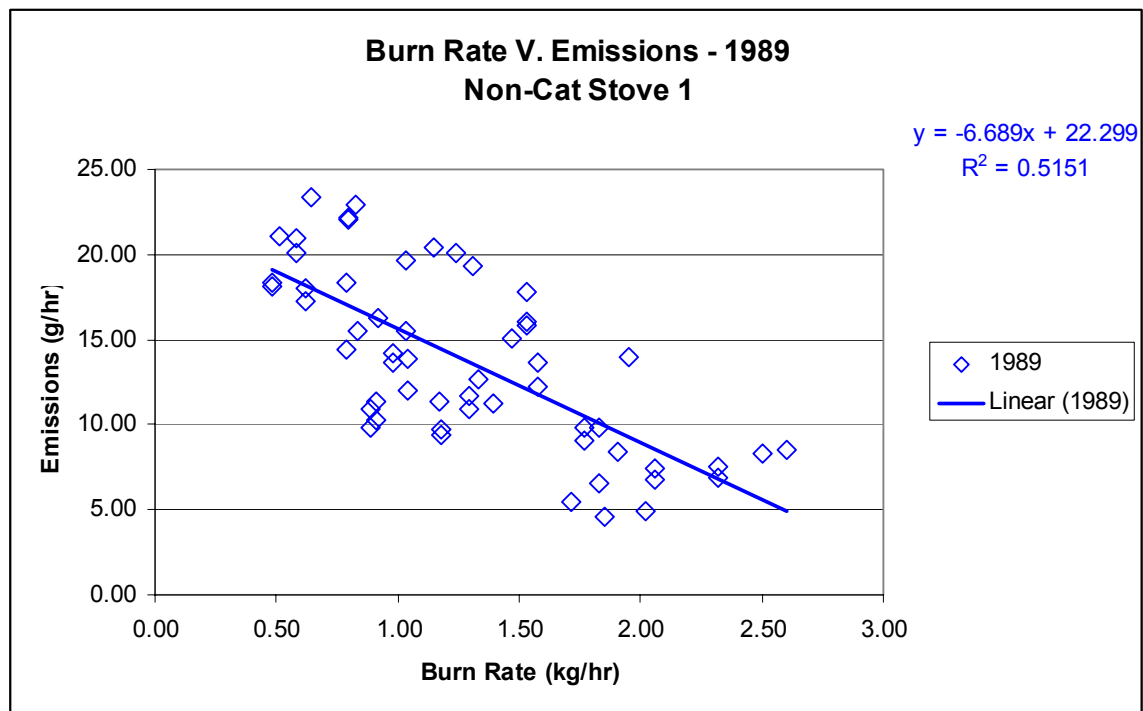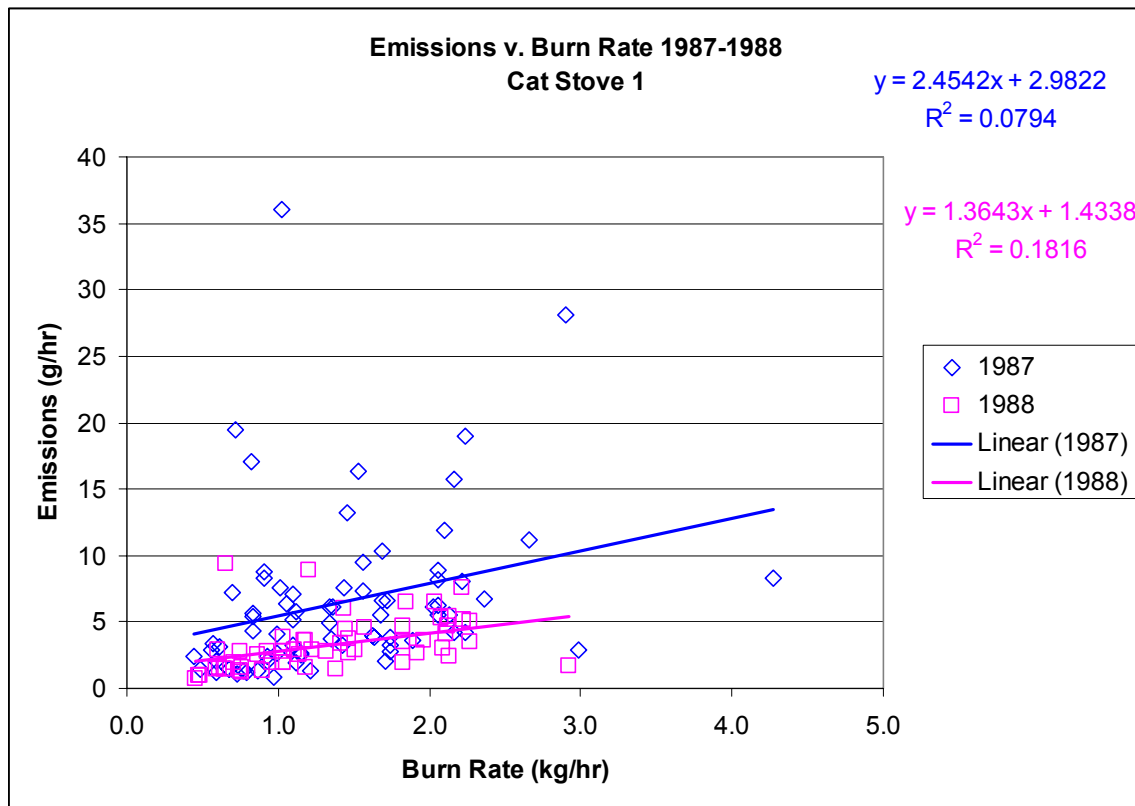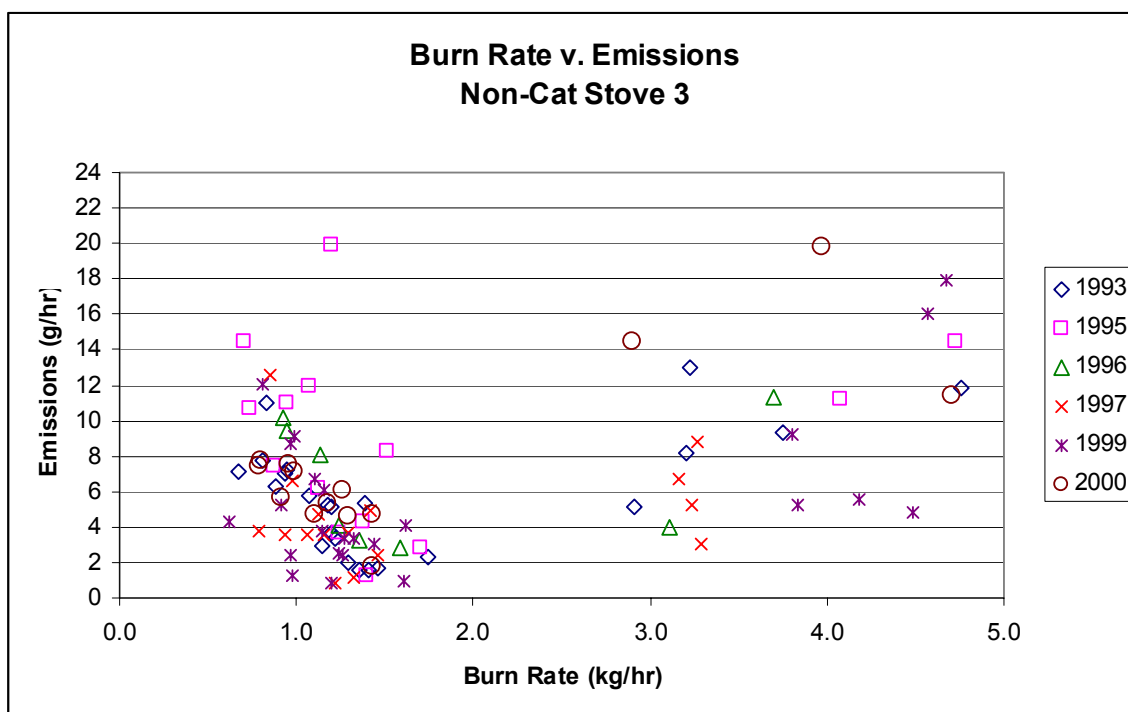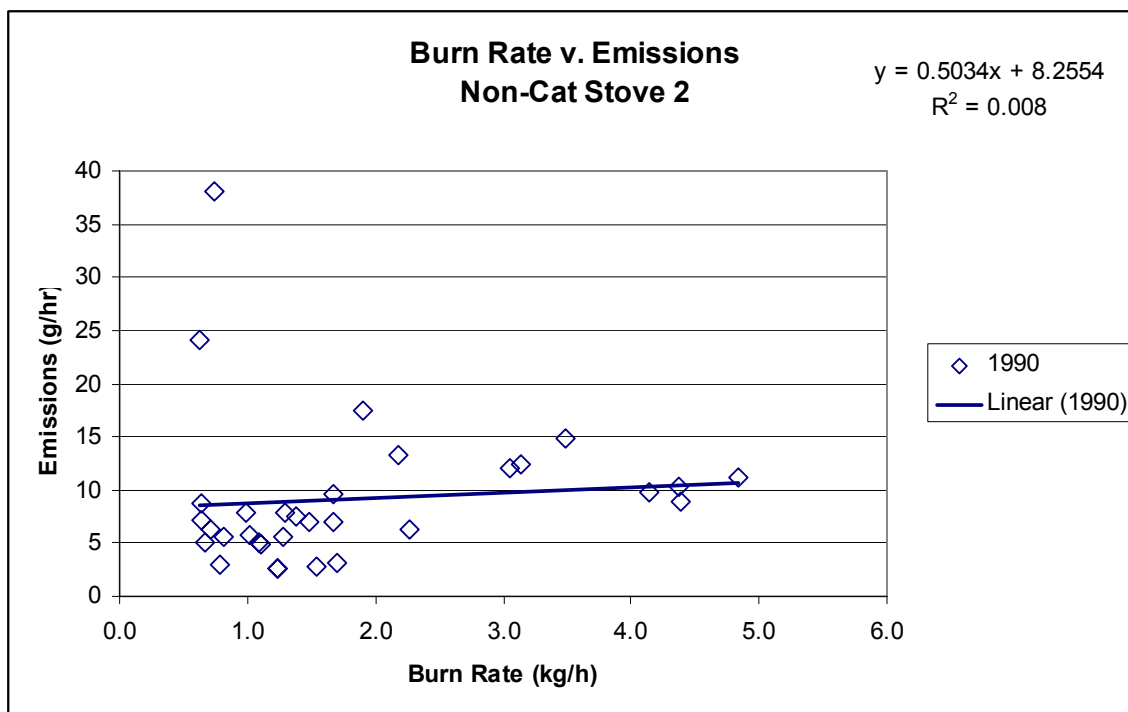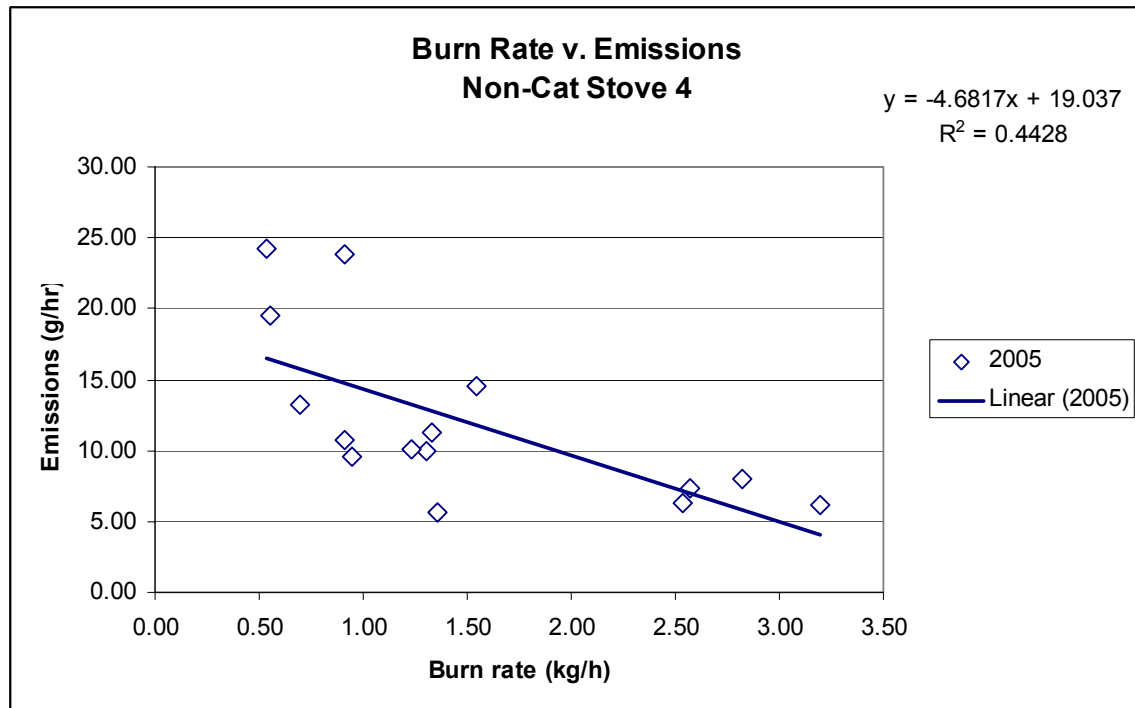
Appendix C - Measurement Uncertainty Analyses for EPA Wood Heater Test Methods

Every numerical physical measurement process is subject to a quantifiable level of uncertainty. This uncertainty is determined in the process of calibrating the measurement instrument. Therefore, the measurement uncertainty for each direct measurement required in a test method can be determined or estimated. When multiple measurements of specific quantities are combined into a final numerical measurement result, the combined uncertainty can be calculated by application of well defined and accepted procedures. These procedures are outlined fully in ISO "Guide to the Expression of Uncertainty in Measurement."

The ISO Guide defines two types of evaluation of measurement uncertainty. Type A is defined as "method of evaluation of uncertainty by statistical analysis of series of observations." This is a quantitative evaluation. Type B is defined as a "method of evaluation of uncertainty by means other than the statistical analysis of series of observations." This describes a more qualitative evaluation. The ISO Guide further elaborates that a "Type B evaluation of standard uncertainty is usually based on scientific judgment using all the relevant information available, which may include:
- previous measurement data,
- experience with, or general knowledge of, the behavior and property of relevant materials and instruments,
- manufacturer's specifications,
- data provided in calibration and other reports, and
- uncertainties assigned to reference data taken from handbooks."

Since there is no such thing as a wood heater test method calibration standard – that is, a wood heater with a "known" emissions rate – the uncertainty calculations and estimates that follow below are based on a hypothetical emissions value and evaluate only the particulate sampling method uncertainty relative to the hypothetical value. Variability in measured emissions due to the combination of the fueling and operation method and general wood burning variability are not components of measurement uncertainty. Measurement uncertainty quantifies only the potential difference between the actual measured result and the "true" value for one specific test run. Variation in results between different test runs includes both the effects of measurement uncertainty and many other variables that influence the actual performance of the product.

An ISO Guide Type A measurement uncertainty analysis is included in this appendix for Method 5G-3 as an example of the way measurement uncertainty is determined using the ISO Guide procedures. Detailed measurement uncertainty calculations are not included for the other methods. However, estimates of the measurement uncertainties for the other 5G methods along with the underlying assumptions used to achieve those estimates are provided. In the case of Method 5H, an estimate of overall method measurement uncertainty[10] is provided. The estimate is based on both ISO Guide Type A and Type B evaluations.

---

[10] This estimate includes both measurement uncertainties that can be analyzed using ISO Guide Type A evaluations and other uncertainties that implicate Type B analyses. The use in this method of one approach for determining stack gas flow rate for setting the proportional sampling rate and another procedure for

EPA Methods 5G-3, 5G-1 and 5G-2

Analysis of the measurement uncertainty for the total particulate emissions in these dilution tunnel test methods requires combining the uncertainty of the following individual measurements.
- Dilution tunnel volumetric flow rate.
- Sampling system volumetric flow rate.
- Filter and sample probe particulate gravimetric catch.

Component Measurement Uncertainties:
- Dilution Tunnel Flow Rate—The dilution tunnel flow is measured by a standard Pitot tube and a differential pressure gauge. The direct measurement is of flow velocity at the operating temperature and pressure. Pitot tube measurements are considered a primary reference method and therefore the primary component of uncertainty in the velocity measurement is that of the differential pressure gauge. Conversion of the velocity measurement to flow rate at standard temperature and pressure conditions requires additional measurements of the tunnel cross sectional area, the temperature and the absolute pressure. An analysis of the overall uncertainty of the tunnel flow rate indicates that the measurements of tunnel cross-sectional area, temperature and pressure have a relatively minor affect. The procedure in this standard has been specified such that the overall uncertainty of this parameter is about ±2 % of the flow rate.
- Sampling Volumetric Flow Rate—The equipment available to measure the sample flow rate has a high precision and can be calibrated such that the measurement uncertainty is ±1 % or better of the actual flow rate. For the purposes of this appendix, a ±1 % uncertainty will be assumed.
- Filter and Probe Particulate Catch—The analytical balance specified in the method has a resolution of ±0.0001g (0.1 mg). For the purposes of measurement uncertainty analysis, we will conservatively assume that the balance resolution equals the measurement uncertainty.

Detailed Measurement Uncertainty (MU) Analysis for Method 5G-3

Weighing Uncertainty
For Method 5G-3, a minimum total of four weighings is needed to determine the net particulate catch (filters and o-rings weighed together and probe weighed separately, before and after the test run). Since each weighing is subject to the same uncertainty, the overall uncertainty of the weighing is:

$$MU_{weighing} = \sqrt{(MU_{Balance})^2 \bullet X}$$

determining stack gas flow rate for the calculation of total emissions is the principle source of measurement error that is assessed using the Type B Analysis. We note, however, that a rigorous analysis of the broader uncertainty implications of this and other Method 5H procedures is outside the scope of this paper.

Page 36 of 43

Where:

MU$_{Balance}$ = Measurement uncertainty of the analytical balance (0.1 mg), and
X = The total number of weight values (4) actually used to calculate emissions. Note: This does not include weight measurements made during pretest or post-test drying, only the final pretest and post-test weights.

$$MU_{weiighing} = \sqrt{0.1^2 \bullet 4} = 0.20$$

For a process that involves two pretest and two post-test weighings, this uncertainty is ±0.20 mg.

An additional component of the filter catch weight uncertainty is the potential for incomplete recovery of the filter material. There is no objective data to determine the magnitude of this factor, but laboratories familiar with the procedures report that they believe it to be less than 0.1 mg. Combining an additional ±0.1 mg with the weighing uncertainty results in an overall weighing uncertainty of ±0.22 mg (0.00022 g).

Combined Measurement Uncertainty
Combined uncertainties are calculated by taking the square root of the sum of squares of the component uncertainties multiplied by a "sensitivity coefficient". The sensitivity coefficient is the partial derivative of the function used to calculate the result with respect to the specific measurement parameter. The general formula (law of propagation of uncertainty) is:

$$uY = \sqrt{((\delta Y/\delta x_1) \times u_1)^2 + ... + ((\delta Y/\delta x_n) \times u_n)^2}$$

Where:
$\delta Y/\delta x_i$ = Partial derivative of the combining formula with respect to individual measurement x_i, and
u_i = the uncertainty associated with that measurement.

The formula to calculate total particulate emissions is:

$$E_T = c_s Q_{std} \theta$$

Where:
c_s = sample filter catch/(sample flow rate x test duration), g/dscf,
Q_{std} = average dilution tunnel flow rate, dscf/min, and
θ = sampling time, minutes.

For the sake of example the following values will be used in an MU analysis of E_T.

Measurement	Measured Value	MU	Units
Sample Filter Catch (F_c)	0.0210	±0.00022	g
Sample Flow Rate (Q_{sample})	0.250	±0.0025	dscfm
Sampling Duration (θ)	180	±0.1	minutes
Tunnel Flow Rate (Q_{std})	150	±3	dscfm

(1) Calculate the MU of c_s:

$$c_s = F_c/(Q_{sample}\,\theta) = 0.0210/(0.25 \times 180) = 0.000467 \qquad g/dscf$$

$$\frac{\delta c_s}{\delta F_c} = \frac{1}{Q_{sample} \bullet \theta} = \frac{1}{0.25 \bullet 180} = 0.0222$$

$$\frac{\delta c_s}{\delta Q_{Sample}} = \frac{-F_c}{Q^2_{sample} \bullet \theta} = \frac{-0.021}{0.25^2 \bullet 180} = -0.001867$$

$$\frac{\delta c_s}{\delta \theta} = \frac{-F_c}{Q_{sample} \bullet \theta^2} = \frac{-0.0210}{0.25 \bullet 180^2} = -0.000003$$

$$MUc_s = \sqrt{(0.00022 \bullet 0.0222)^2 + (0.0025 \bullet -0.001867)^2 + (0.1 \bullet -0.000003)^2}$$

$$MUc_s = \mathbf{0.000007 \ g/dscf}$$

Thus, c_s would be 0.467 mg/dscf ± 0.007 mg/dscf at a 95 % confidence level.

(2) Calculate E_T and MU_{ET}

$$E_T = c_s Q_{std}\,\theta = .000467 \times 150 \times 180 = 12.60 \ g$$

$$\frac{\delta E_T}{\delta c_s} = Q_{std} \bullet \theta = 150 \bullet 180 = 27,000$$

$$\frac{\delta E_T}{\delta Q_{std}} = c_s \bullet \theta = 0.000467 \bullet 180 = 0.084060$$

$$\frac{\delta E_T}{\delta \theta} = c_s \bullet Q_{std} = 0.000467 \bullet 150 = 0.070050$$

$$MU_{ET} = \sqrt{(27,000 \bullet 0.000007)^2 + (0.084060 \bullet 3)^2 + (0.070050 \bullet 0.1)^2} = \textbf{0.315 g}$$

Thus the result in this example would be:

***ET* = 12.60 g ± 0.32 g at a 95 % confidence level or stated as MU % = ±2.5%.**

Conclusion—This example, which is representative of the measurement method as it is currently applied to woodstoves under the EPA NSPS, indicates that the uncertainty related to the dilution tunnel flow rate measurement and filter catch weights are the primary components of the overall uncertainty of the result. Tunnel flow rates may be much higher than necessary to capture all emissions. This can influence the uncertainty of the test method. For example, increasing the tunnel flow rate from 150 to 600 scfm in the above example increases the MU of the result to ±0.8 g or about ±6.5 % of the measured emissions. Therefore, keeping tunnel flow rates near the minimum necessary to reliably capture the exhaust stream, while keeping the tunnel velocity at a level that can be accurately measured during tests will minimize the uncertainty of the measurement.

Measurement Uncertainty Estimate for EPA Method 5G-1

The major difference between EPA Method 5G-3 and 5G-1 is that 5G-1 uses solvent recovery to determine the particulate weight collected in the probe and front half of the front filter holder. The requirement to weigh a solvent blank results in at least one additional weighing before and after the test when compared to Method 5G-3. It is also common practice to weigh the 100mm filters separately which also results in another weighing before and after the test. These additional weighings add to uncertainty. There is also the likelihood of more substantial fiber loss while recovering adhered filter fibers from the filter gaskets with the 100 mm filters and the possibility of some loss of particulate from the probe during the solvent recovery process.

If the same balance uncertainty of 0.1 mg for six total weighings is combined with a conservative estimate of 0.1% of total catch for the combination of fiber loss from the gaskets and particulate loss during solvent recovery is used, the resultant uncertainty for the filter and probe catch weight is significantly higher than with Method 5G-3.

However, the sample flow rate, and therefore the probe and filter catch, with Method 5G-1 is typically about twice as high as Method 5G-3 and this difference offsets much of the impact of the increase in weighing uncertainty on the overall method measurement uncertainty. Following the same analysis track as shown for Method 5G-3 above, the Method 5G-1 MU is about 10 - 15% higher than the MU for 5G-3. This MU% range of ±2.7 to 3% is still quite good.

Measurement Uncertainty Estimate for EPA Method 5G-2

Using the same logic applied to the Method 5G-1 measurement uncertainty estimate, the primary difference between 5G-1 and 5G-2 is in the number of filter train components subjected to solvent recovery and the total number of weighings before and after each test run. After consultation with a highly experienced 5G-2/5H test technician, it was determined that there are typically seven individual weight measurements before and after each test run. These include four beakers used for solvent recovery, a beaker for the solvent blank and the front and rear filters. These additional weighings contribute to additional uncertainty.

It would not be unreasonable to assume that some additional fraction of particulate matter could be lost during solvent recovery due to all the additional components that are involved in the recovery. But, even if we continue to assume that a very conservative 0.1% of the total particulate catch is lost from filter gaskets and solvent recovery, the weighing uncertainty increases by about 40% over Method 5G-1. However, since the total catch is expected to be somewhat higher with 5G-2, the larger catch mitigates some of the impact of the increased weighing uncertainty on the overall method measurement uncertainty. We would expect that the measurement uncertainty for Method 5G-2 is in the range of 15 -20% higher than Method 5G-1. The MU% is in the range of ±3 – 3.5% which is also quite good.

Overall Measurement Uncertainty Estimate for EPA Method 5H

The same particulate sampling equipment is used for both Method 5H and Method 5G-2, but the methods are not close to the same in terms of overall measurement uncertainty. If we first examine the sampling protocols for the two methods, we find there are notable differences. These include:

- The particulate concentrations in the 5H sampling stream are much higher than 5G-2 (undiluted stack gas in 5H versus diluted gas in the tunnel with 5G-2).
- Lower sampling rates are typically used for 5H to prevent filter plugging (0.1 – 0.3 cfm for 5H and 0.5 cfm or higher for 5G-2).
- Total particulate catches are higher (by an order of magnitude or more) for 5H compared to 5G-2.

However, despite these differences, the particulate catch measurement uncertainty for Method 5H is not appreciably higher than that for 5G-2. This is because the impacts of the differences noted above cut to some extent in different directions, e.g., the added uncertainty contributed by the low sampling rate is somewhat mitigated by the much higher total particulate catch. Since the measurement uncertainty in the sample flow rate

and the total particulate catch parameters are reasonably quantifiable, we would estimate their combined contribution to the overall measurement uncertainty of Method 5H using an ISO Guide Type A analysis to be no more than a few percentage points. At this level, the sampling-related measurement uncertainty is only a secondary contributor to the overall measurement uncertainty of Method 5H.

The primary contributor to the large increase in overall measurement uncertainty associated with Method 5H when compared to Method 5G-2 (or the other dilution tunnel methods) is the determination of the stack gas flow rate over the course of the test. The associated complexities with this issue make quantitative assessments of the contribution to measurement uncertainty unworkable using ISO Guide Type A evaluations. The more qualitative Type B evaluation is appropriate in this case.

Factors Contributing to Stack Flow Uncertainty

The stack flow must be determined for purposes of knowing where to set the sample flow rate in order to maintain sample flow proportionality over the test run. The procedure specified in Method 5H is a tracer gas method. This method has been generally replaced by adding a dilution tunnel to the system and determining stack flow using the ratio of tunnel CO_2 to stack CO_2 multiplied by the easily measured tunnel flow. At each sampling interval, the sample flow is adjusted based on this stack flow determination. There is additional uncertainty in this procedure if a single gas analyzer is used to determine both the stack and tunnel CO_2 concentrations since the accuracy of these instruments is generally 1% of full scale and the tunnel CO_2 readings are quite low. Some of this can be mitigated by using two analyzers or multi-scale analyzers where the accuracy can better match the concentrations.

The 5H procedures then go on to calculate stack gas flow for the purpose of determining emissions using a mass balance algorithm where carbon from the fuel is balanced against the carbon measured in the stack (CO_2 and CO) and an assumed hydrocarbon value to make the balance work. The stack flow calculated by this means often results in very different incremental stack flow values than the tracer gas methods. The impact of this difference is that the proportionality of the sample gas flow rate to the stack gas flow rate can be much different between the two procedures. Additionally, there is a wider range in variation from proportional sample flow rate allowed in this method than for the dilution tunnel methods. The allowable range is a necessity due to the fact that frequent stack gas flow calculations and sample flow rate adjustments are needed due to the instability in the actual stack gas flow. The impact of proportionality ranges is hard to asses but contribution to method uncertainty is definitely real. Finally, the total stack gas flow over the entire test run is also impacted by the mass balance procedure and this obviously directly impacts the particulate emission determination. The uncertainty of the carbon balance stack flow rate calculation is impacted by the unknown uncertainties of the fuel carbon content, the assumed hydrocarbon concentration and the measurement uncertainty of stack CO and CO_2 measurements.

Overall measurement uncertainty for 5H is commonly estimated in the ±20 – 30% range by those with considerable experience with the method. The increment of measurement uncertainty that can be analyzed using ISO Guide Type A procedures accounts for only a

few percentage points of this range. The remainder (~17 – 27%) can be supported by a qualitative (Type B) analysis using reasonable estimates of the measurement uncertainties associated with the stack flow procedures in Method 5H. This estimate corresponds to the magnitude of the typical difference between stack flow rates determined by the tracer gas and carbon balance procedures for an individual test run.

REFERENCES

1. 40 CFR, Part 60, Subpart AAA – Standards of Performance for New Residential Wood Heaters.

2. EPA Woodstove NSPS Proficiency Test Results – Supplied by Michael Toney – EPA OAQPS, Emissions Measurement Branch.

3. Guide to the Expression of Uncertainty in Measurements, First Edition 1995, International Organization for Standards.

4. ASTM E177-2008 "Standard Practice for Use of the Terms Precision and Bias in ASTM Test Methods", ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959.

5. ASTM E691-2009 – "Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method", ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959.

6. CSA B415.1-2010 - Performance Testing of Solid-Fuel-Burning Heating Appliances, CSA International, 5060 Spectrum Way, Suite 100, Mississauga, Ontario, Canada L4W 5N6, March, 2010.

7. ASTM E2515 – 2007 – "Standard Test Method for Determination of Particulate Matter Emissions Collected by a Dilution Tunnel", ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959.

8. ASTM E2618 – 2009 – "Standard Test Method for Measurement of Particulate Emissions and Heating Efficiency of Outdoor Solid Fuel-Fired Hydronic Heating Appliances", ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959.

9. Mark's Standard Handbook for Mechanical Engineers, Ninth Edition, Eugene A Avallone, Theodore Baumeister III, McGraw Hill Book Co., 1987.

Commenter Name: Mary Sullivan Douglas
Commenter Affiliation: National Association of Clean Air Agencies (NACAA)
Document Control Number: EPA-HQ-OAR-2002-0058-3525-A1
Comment Excerpt Number: 11

Comment: The first step in any scientifically sound measurement process is to ensure that the procedures employed are sufficiently precise to determine meaningful differences. In response to questions from industry as to whether they should extend sampling periods to ensure more precise results, EPA advised them that they did not need to and that the agency would address it the final rulemaking. EPA defines the method detection limit as, "the minimum concentration of a substance that can be measured and reported with 99% confidence that the analyte concentration is greater than zero and is determined from analysis of a sample in a given matrix containing the analyte." Where the "adjusted" average emissions of the top 12 percent is "near" the method detection level, EPA now proposes[10] to increase the calculated average so that the floor is not less than 300 percent of the detection level. To justify this increase EPA observes that when measurements are near the detection level the measurement uncertainty can be as high as (+/-) 40 percent, while such uncertainty is reduced to (+/-) 15 percent if the measured value is three times (300 percent) the detection level. However, since such measurement uncertainties are necessarily part of the overall variability determined in step one of EPA's procedure, there is no need or basis to substitute this arbitrary figure for the actual emission data that the statute requires be used. Additionally, it also makes no technical sense to introduce a known error of 300 percent in the MACT floor in order to avoid a possible error of 25 percent[11] in any individual measurement. This step constitutes yet one more bias in favor of allowing higher levels of HAP emissions. In this rulemaking EPA proposes to compound this error by "adjusting" the detection level reported by the laboratory in accordance with established protocols, even where EPA has no information that the detection levels reported by the laboratory are incorrect.

[Footnotes]

(10) EPA employed this technique in the cement kiln New Source Performance Standard rule.

(11) This is the difference between the potential error at the detection level and that at three times the detection level.

Response: See the response to comment EPA-HQ-OAR-2002-0058-3511-A1, excerpt 14.

We disagree with the commenter assertion that EPA, knowing many results would be at or below detection limits if it only required sample periods of one hour per run, responded to industry that the sources need not extend sample periods to ensure more precise results. In fact, enclosure 1 to the ICR letter requesting testing stated the methods and sampling times or volumes. For example, for sampling for metals, it was stated that EPA Method 29 be used and to collect a minimum volume of 4.0 cubic meters or have a minimum sample time of 4 hours per run. For dioxins/furans, enclosure 1 specified that EPA Method 23 be used and to collect a minimum volume of 2.5 cubic meters or have a minimum sample time of 4 hours per run.

Comment: Insetting the Boiler MACT standards,EPA has acknowledged that the emission limit should not be set below the capability of the applicable test method. However, EPA did not use the widely accepted definition of method detection limit, which is based on the capabilities of multiple commercial laboratories to analyze a sample and identify the presence of a chemical above the "noise" level. In its place, EPA coined a new term, "representative method detection limit" (RDL) to define a measurement method detection limit which is based on the laboratory detection limits reported for the tests with the lowest emissions.This erroneous methodology resulted in estimating D/F detection limits that are over 100 times lower than those regularly achieved by commercial laboratories, based on an analysis documented in previous comments by AF&PA and the National Council for Air and Stream Improvement (NCASI).20

The detection limit of an analytical method is commonly defined as the lowest concentration that can be distinguished from replicate blanks.The quantitation limit of a method is defined as the smallest concentration of the substance which can be measured with an acceptable level of uncertainty. Detection limits and quantitation limits are defined in a scientific, non-arbitrary manner in various widely-published peer-reviewed consensus guidelines 21 and EPA documents. Quantitation limits of test methods have great significance when measuring very low concentrations of pollutants. In practice, reported values below the method's quantitation limit should not be treated as real values.

The majority of a federal advisory committee on method detection and quantitation limits recommended that the quantitation limit of a test method should be based on the 95th percentile of what is being achieved by the commercial laboratories.22 Using this approach and rounding up the quantitation limit to a single digit,as EPA has done for floor setting, would result in a D/FTEQ quantitation limit of 0.2 ng/dscm based on analytical procedures alone. However, using the results of the ASME ReMAP study,23 which evaluated the precision of Method 23D/F stack testing measurements, a D/F TEQ quantitation limit of 0.27 ng/dscm is derived, which addresses the uncertainty of all of the test method components.

[Footnote 20: The National Council for Air and Stream Improvement is an independent, non-profit research institute that focuses on environmental topics of interest to the forest products industry. Established in 1943, NCASI is recognized as the leading source of reliable data on environmental issues affecting this industry.]

[Footnote 21: For example, Keith, L.H.et al.,Principles of Environmental Analysis, *Anal.Chem*.1983,*55*,2210-2218.]

[Footnote 22: Report of the federal advisory committee on detection and quantitation approaches and uses in Clean Water Act Programs submitted to the Environmental Protection Agency (EPA). December 2007.]

[Footnote 23: Lanier, W.S. and Hendrix, C.D. Final Report--Reference Method Accuracy and Precision (ReMAP): Phase 1 Precision of Manual Stack Emission Measurements, ASME International, Washington, D.C. (2001).]

Response: For a response to the comment that Tesoro's boiler SG1102 test data are not representative of other non-continental liquid units due to low-load operation and site-specific fuel, please see comment EPA-HQ-OAR-2002-0058-3677-A2, excerpt 64. EPA has incorporated additional CO emissions testing data for the top performer into the recalculation of the 99% UPL for the subcategory. The variability between the previously reported and new emissions testing data resulted in an increased CO emission limitation for boilers designed to combust liquid fuel located in non-continental states or territories.

3C. MACT Floor Methodology: Non-Detect Values

Commenter Name: James Pew
Commenter Affiliation: Earthjustice, Clean Air Council, Partnership for Policy Integrity
Document Control Number: EPA-HQ-OAR-2002-0058-3511-A1
Comment Excerpt Number: 14

Comment: EPA exacerbated these problems with the approach to "non-detect" test results that it took in 2010. That approach continues to be unlawful and arbitrary for all the reasons given in the 2010 comments, which are incorporated by reference as if fully stated herein and reiterated with respect to the agency's 2011 final rule and reconsideration proposal. 2010 Comments at 22.

Now, EPA further exacerbates the other flaws in its variability analysis by arbitrarily assuming that individual sources' actual emission level are three times <u>higher</u> (*i.e.*, worse) than an arbitrarily selected "method detectionlevel"(MDL)whenever their emissions fell <u>below</u> this level. That bizarre assumption follows a series of decisions that were unlawful, arbitrary, or both. First, EPA itself encouraged sources to report tests below method detection levels by advising them during the data collection phase that they did not need to use specific (or even adequate) criteria for testing and assuring them that the agency would address the absence of precise data in its rulemaking. As a result of EPA's stated willingness to accept data that were not precise, many sources used testing methods that were not precise and that yielded non-detect results at emission levels that could have been precisely measured had these sources chosen to use more precise testing methods – which are available and have been available for many years. Second, EPA now claims that any test below the mean detection level for all the alleged best sources in every subcategory – a level the agency now dubs the "representative detection level" or "RDL" – is a non-detect, whether that test really is a non-detect or not. Third, for all of the alleged non-detect results, EPA simply fills in a fictional replacement value equal to 300% of the RDL, unless that number would be lower than the floor. Nowhere does the agency offer any explanation for its apparent assumption that the 99% UPL it applies to each source's emission test results does not already fully account for that source's variability. Further EPA uses this 300% multiplier even though it does not and cannot provide any rational basis for assuming that all test results below the RDL actually reflect emission levels 300% higher than the RDL (even assuming *arguendo* that that they are higher than the RDL at all). All the agency offers on that subject is a conclusory assertion that multiplying the RDL by three approximates a 99% UPL for a data set of seven or more values, without either supporting that claim or explaining why it is even relevant. For these reasons alone, EPA's floors do not reflect the best sources' actual emission levels and are unlawful and arbitrary.Moreover, it is arbitrary – and amounts to a sabotage of the rule – for EPA to encourage sources to submit imprecise emissions data and then

use the resulting imprecision in the data it receives as an excuse to grossly inflate the emission standards.

Response: The commenter is incorrect that EPA's test method specifications during the ICR data collection encouraged the collection of data below the detection level. During the data collection request, we specified sampling times and/or volumes where appropriate. These times/volumes were calculated to provide reasonable assurance of detecting compounds that were present. EPA manual testing methods are designed to be conducted over an hour of sampling, but may be conducted with extended testing times to collect additional sample volume and provide for increased measurement sensitivity. Our guidance to ICR respondents for reporting pollutant emissions used to support floor development has been to require them to provide test-specific method detection limit (MDL) values in the reports. Also, in accordance with our guidance, source owners are to identify emissions data which were measured below the MDL and report those values as equal to the MDL determined for that test. This is done to provide for minimum pollutant concentrations when evaluating the emission floor, otherwise the measured concentration values as reported in the testing were used in calculating the floor.

Not all test reports that include instrumental test method data included test-specific MDL values. In cases where the report does not include test-specific MDL data measured with an instrumental test method, the test-specific MDL values were determined using the reported calibration span values. The EPA accounted for the effect of measurement imprecision in calculating a floor using a database that includes reported MDL data by first defining an MDL value that is representative of the data to be used in establishing the floor or emissions limit. This value is termed the representative method detection limit (RDL). The second step in the process was to calculate three times the RDL and compare that value to the calculated floor or emissions limit. The value of 3* RDL is necessary to gauge the performance test methods' level of quantitation, or that point at which the test method begins to return values within expected levels of confidence. The EPA recognizes that values between the method detection limit and the level of quantitation have more uncertainty than values at or above the level of quantitation, therefore we make this determination to provide a value to the floor setting decision process that describes the level of quantitation for the compliance determination method. If 3*RDL was less than the calculated floor or emissions limit calculated from the upper prediction limit (UPL), we concluded that measurement variability was adequately addressed and the calculated floor or emissions limit was not adjusted. If, on the other hand, the value equal to 3*RDL was greater than the floor value or emission limit, we concluded that the calculated floor or emission limit does not account entirely for measurement uncertainty, and the value equal to 3*RDL was substituted for the adjusted floor or emissions limit (i.e., increasing the floor to be equal to 3*RDL). This adjusted value ensures measurement uncertainty is adequately addressed in the floor or the emissions limit and that the test method used to determine compliance with the emission standard will be appropriate for quantitative determination of emissions concentrations at, and above, the level of the emission standard.

Originally, the floor was calculated by multiplying individual tests by a factor of three when the result was below the RDL. This methodology has been corrected in the final rule to the standard methodology described above.

See the memo from Peter Westlin to Docket EPA-HQ-OAR-2002-0058, April 10, 2012, for further information on floor determination for instrumental test methods.